

1993

The role of emotion in rationality: Limiting the search for evidence.

John Stewart. McKay
University of Windsor

Follow this and additional works at: <http://scholar.uwindsor.ca/etd>

Recommended Citation

McKay, John Stewart., "The role of emotion in rationality: Limiting the search for evidence." (1993). *Electronic Theses and Dissertations*. Paper 4239.

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file - Votre référence

Our file - Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

The Role of Emotion in Rationality:
Limiting the Search for Evidence

By

John McKay

A Thesis
Submitted to the Faculty of Graduate Studies and Research
Through the Department of Philosophy
in Partial Fulfillment of the Requirements for
the Degree of Master of Arts at the
University of Windsor

Windsor, Ontario, Canada

1993

(c) John Stewart McKay



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file *Votre référence*

Our file *Notre référence*

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-83021-2

Canada

Name

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

Philosophy

SUBJECT TERM

0422

U·M·I

SUBJECT CODE

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS

Architecture	0729
Art History	0377
Cinema	0900
Dance	0378
Fine Arts	0357
Information Science	0723
Journalism	0391
Library Science	0399
Mass Communications	0708
Music	0413
Speech Communication	0459
Theater	0465

EDUCATION

General	0515
Administration	0514
Adult and Continuing	0516
Agricultural	0517
Art	0273
Bilingual and Multicultural	0282
Business	0688
Community College	0275
Curriculum and Instruction	0727
Early Childhood	0518
Elementary	0524
Finance	0277
Guidance and Counseling	0519
Health	0680
Higher	0745
History of	0520
Home Economics	0278
Industrial	0521
Language and Literature	0279
Mathematics	0280
Music	0522
Philosophy of	0998
Physical	0523

Psychology	0525
Reading	0535
Religious	0527
Sciences	0714
Secondary	0533
Social Sciences	0534
Sociology of	0340
Special	0529
Teacher Training	0530
Technology	0710
Tests and Measurements	0288
Vocational	0747

LANGUAGE, LITERATURE AND LINGUISTICS

Language	
General	0679
Ancient	0289
Linguistics	0290
Modern	0291
Literature	
General	0401
Classical	0294
Comparative	0295
Medieval	0297
Modern	0298
African	0316
American	0591
Asian	0305
Canadian (English)	0352
Canadian (French)	0755
English	0593
Germanic	0311
Latin American	0312
Middle Eastern	0315
Romance	0313
Slavic and East European	0314

PHILOSOPHY, RELIGION AND THEOLOGY

Philosophy	0422
Religion	
General	0318
Biblical Studies	0321
Clergy	0319
History of	0320
Philosophy of	0322
Theology	0469

SOCIAL SCIENCES

American Studies	0323
Anthropology	
Archaeology	0324
Cultural	0326
Physical	0327
Business Administration	
General	0310
Accounting	0272
Banking	0770
Management	0454
Marketing	0338
Canadian Studies	0385
Economics	
General	0501
Agricultural	0503
Commerce-Business	0505
Finance	0508
History	0509
Labor	0510
Theory	0511
Folklore	0358
Geography	0366
Gerontology	0351
History	
General	0578

Ancient	0579
Medieval	0581
Modern	0582
Block	0328
African	0331
Asia, Australia and Oceania	0332
Canadian	0334
European	0335
Latin American	0336
Middle Eastern	0333
United States	0337
History of Science	0585
Law	0398
Political Science	
General	0615
International Law and Relations	0616
Public Administration	0617
Recreation	0814
Social Work	0452
Sociology	
General	0626
Criminology and Penology	0627
Demography	0938
Ethnic and Racial Studies	0631
Individual and Family Studies	0628
Industrial and Labor Relations	0629
Public and Social Welfare	0630
Social Structure and Development	0700
Theory and Methods	0344
Transportation	0709
Urban and Regional Planning	0999
Women's Studies	0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES

Agriculture	
General	0473
Agronomy	0285
Animal Culture and Nutrition	0475
Animal Pathology	0476
Food Science and Technology	0359
Forestry and Wildlife	0478
Plant Culture	0479
Plant Pathology	0480
Plant Physiology	0817
Range Management	0777
Wood Technology	0746
Biology	
General	0306
Anatomy	0287
Biostatistics	0308
Botany	0309
Cell	0379
Ecology	0329
Entomology	0353
Genetics	0369
Limnology	0793
Microbiology	0410
Molecular	0307
Neuroscience	0317
Oceanography	0416
Physiology	0433
Radiation	0821
Veterinary Science	0778
Zoology	0472
Biophysics	
General	0786
Medical	0760

EARTH SCIENCES

Biogeochemistry	0425
Geochemistry	0996

Geodesy	0370
Geology	0372
Geophysics	0373
Hydrology	0388
Mineralogy	0411
Paleobotany	0345
Paleoecology	0426
Paleontology	0418
Paleozoology	0985
Palynology	0427
Physical Geography	0368
Physical Oceanography	0415

HEALTH AND ENVIRONMENTAL SCIENCES

Environmental Sciences	0768
Health Sciences	
General	0566
Audiology	0300
Chemotherapy	0992
Dentistry	0567
Education	0350
Hospital Management	0769
Human Development	0758
Immunology	0982
Medicine and Surgery	0564
Mental Health	0347
Nursing	0569
Nutrition	0570
Obstetrics and Gynecology	0380
Occupational Health and Therapy	0354
Ophthalmology	0381
Pathology	0571
Pharmacology	0419
Pharmacy	0572
Physical Therapy	0382
Public Health	0573
Radiology	0574
Recreation	0575

Speech Pathology	0460
Toxicology	0383
Home Economics	0386

PHYSICAL SCIENCES

Pure Sciences

Chemistry	
General	0485
Agricultural	0749
Analytical	0486
Biochemistry	0487
Inorganic	0488
Nuclear	0738
Organic	0490
Pharmaceutical	0491
Physical	0494
Polymer	0495
Radiation	0754
Mathematics	0405
Physics	
General	0605
Acoustics	0986
Astronomy and Astrophysics	0606
Atmospheric Science	0608
Atomic	0748
Electronics and Electricity	0607
Elementary Particles and High Energy	0798
Fluid and Plasma	0759
Molecular	0609
Nuclear	0610
Optics	0752
Radiation	0756
Solid State	0611
Statistics	0463

Applied Sciences

Applied Mechanics	0346
Computer Science	0984

Engineering	
General	0537
Aerospace	0538
Agricultural	0539
Automotive	0540
Biomedical	0541
Chemical	0542
Civil	0543
Electronics and Electrical	0544
Heat and Thermodynamics	0348
Hydraulic	0545
Industrial	0546
Marine	0547
Materials Science	0794
Mechanical	0548
Metallurgy	0743
Mining	0551
Nuclear	0552
Packaging	0549
Petroleum	0765
Sanitary and Municipal	0554
System Science	0790
Geotechnology	0428
Operations Research	0796
Plastics Technology	0795
Textile Technology	0994

PSYCHOLOGY

General	0621
Behavioral	0384
Clinical	0622
Developmental	0620
Experimental	0623
Industrial	0624
Personality	0625
Physiological	0989
Psychobiology	0349
Psychometrics	0632
Social	0451



Abstract

In this thesis I consider a key problem in the theory of rationality, namely that as limited knowers we must set limits on the scope of our considerations.

In Chapter I outline the problem: the need to draw certain kinds of cognitive limits in order to be rational and the special role that emotions play in that process. In Chapter II, I outline a theory of rationality. I do this by considering a number of views on rationality and epistemology and by pointing out their strengths and flaws in order to arrive at my own account of rationality. My own view emphasizes the problem of our limited capacity as "knowers" and how we can accept that and still pursue truth and knowledge. Central to this is the need to limit the search for evidence; we must accept our reasoning ability as limited and come to closure. In Chapter III I consider some prominent theories of emotion in order to understand what role they might contribute to a theory of rationality. In Chapter IV I consider de Sousa's arguments for the conclusion that emotions have an essential role in creating salience, and in so doing help us overcome (i) cognitive binds of indecision (Buridan's Ass) and (ii) the need to exclude possibly relevant evidence (the frame problem). I argue, against de Sousa, that emotion is not required to free us from a deterministic rationality yet agree with his speculation that emotion might help to solve cognitive binds by mimicking cognitive encapsulation.

Acknowledgements

I would like to thank Robert C. Pinto for guiding me through this project and my parents for their financial support throughout my education.

Contents

Abstract iii

Acknowledgements iv

Chapter I: INTRODUCTION 1

Outline 5

Chapter II: A THEORY OF RATIONALITY 7

1 Introduction 7

2 Naturalizing Rationality 8

3 A Note Regarding Irrationality 9

4 Defining Questions 12

5 Cognitive versus Strategic Rationality: de Sousa 14

6 Practical and Evaluative Rationality: Rescher 23

7 The Predicament of Reason 25

8 Evaluative Rationality 29

9 Rationality and Principles 30

10 The Interdependence of Different Types of Reason 32

11 James' View of Reason, Belief and Truth 34

12 Conservatism, Coherence and the Fabric of Belief 41

13 Harman and Naturalized Rationality 46

14 The Open Ended Discourse of Discourses 56

15 Meta-Cognition and Our Lack of Control 59

16 Conclusion 62

Chapter III: PHILOSOPHICAL THEORIES OF EMOTION 64

1 Introduction 64

2 Emotion and Rationality 64

3 Plato 66

4 Hume 67

5 Theories of Emotion 72

6 The Structure of Emotion: Causal Genesis 77

7 The Structure of Emotion: Intentional Adequacy 79

8 Conclusion 88

Chapter IV: THE ROLE OF EMOTIONS IN RATIONAL COGNITION 89

1 Introduction 89

2 The Determinism of Reasons and Buridan's Ass 90

3 Emotion and Salience 99

4 Rationality and the Frame Problem 102

5 Emotion and Encapsulation 105

Chapter V: CONCLUSION 109

APPENDIX I: Why Reason? 111

APPENDIX II: CONTEXT AND METHODOLOGY 114

1 Introduction 114

2 Mind Body 115

3 Free Will, Determinism, Compatibilism 116

4 Modes of Description 119

BIBLIOGRAPHY 125

VITA AUCTORIS 128

I INTRODUCTION

In this thesis I want to consider what I think is a key problem in the theory of rationality, namely that as limited knowers we must set limits on our scope of considerations. In our cognitive functioning we frequently must gather information either by searching our memory or by looking outside ourselves. I want to argue that due to our epistemological position it is difficult for us to say exactly when we must stop the information gathering and reach at least tentative conclusions. This can occur either on a reflective conscious level where we have recourse to a normative theory of thought or rationality, or it can be seen as occurring at a pre-conscious level in virtue of certain aspects of our functional cognitive architecture. I will explore both these possibilities. In some sense what I'm interested in is how we can know something without knowing everything.

Another way of understanding this is as a question of the nature of relevance. Unlike us, an omniscient being, were it to need to reason at all (which it doesn't), would know exactly what is relevant in determining the truth of anything. All the connections would be laid out before it. Since we do not have this capacity we must proceed inductively. Now, in inductive reasoning relevance is a contingency; when we are trying to determine the truth of something there is no determinate a priori process by which we can say exactly what will be relevant to achieving that. We may, for instance, think that we have established the truth of something and then upon gathering some new piece of information decide that we were mistaken in that previous judgment. Very often there is no way that we could have determined that we needed that specific piece of information

until we discovered that we did in fact need it. I don't think that this just applies to hypothetical reasoning, but rather hypothetical reasoning is a more clearly defined process that is relevantly similar, for my purposes, to investigation generally.

These considerations lead to a pragmatic and naturalized theory of rationality where the chief requirement is that it work. This means that a theory of rationality must accept the realities of human cognition when it attempts to describe how we should think. Kant sums up everything that can go wrong when we try to address the nature of rationality when he writes:

Logic can have no empirical part - that is, no part in which the universal and necessary laws of thinking are based on grounds taken from experience. Otherwise it would not be logic - that is, it would not be a canon for understanding and reason, valid for all thinking and capable of demonstration. (Kant,55)

Logic, or any other a priori science, is quite limited in helping us understand and guide "understanding and reason". This is not due to the problems or weaknesses of logic, but rather with the demands of a theory of rationality. Rationality must deal with the problem of how limited minds cope with an infinite (or at least indefinitely large) universe. God's all-encompassing view needs no principles to guide it in the understanding of the world; it simply knows all in one glance. This, however, is not our situation.

What this situation implies is that we are always lacking in our attempts to understand and account for the world. Nicholas Rescher refers to this situation as the "predicament of reason"; reason demands of us that we move forward in seeking truth and come to some conclusions. We are, nonetheless, always faced with the possibility that our conclusions are

wrong because in meeting the first injunction we were forced to limit our purview of reasons somewhat arbitrarily. This means that we may discover upon further consideration that we were wrong despite the fact that everything up til then made it look like we were right.

The sceptic understands this point and argues that truth is not attainable. I argue that armed with this information we can undertake an epistemology that appreciates that our knowledge or belief structures are inherently open-ended systems where what is lacking, and our decisions about where to curtail investigation, are often just as important as what we include in our beliefs. A normative theory for thought will have to take into account the contingencies of human thought.

Furthermore, implicit in Kant's view above of the nature of rationality, I believe, is the view that rationality is quite independent of other aspects of mind, especially that aspect of mind (and corporeality) we call emotion. I argue here that emotion has an important role to play in rational cognition.

The relation of emotion to rationality can be viewed as an aspect of the investigation that we would make into the nature of rationality once we have given up the notion that logic as an *a priori* science is to be the canonical guide (both normatively and descriptively) with which we approach rationality. When we move beyond this notion, as Harman does in *Change In View*, we need to look, independently of logic as such, for what we would want to call a rational functioning of the human mind.

I develop here a notion of rationality as optimization or intelligent procedure; doing things as best they can be done. The objectivity for the notion of "best" is derived from judgments about how well things achieve

their formal objects. The formal object of any intentional state is that toward which the state appropriately aims. For example, the formal object of belief is truth. Exactly how well any *procedure* might achieve its formal object is a question to be determined by empirical and conceptual research, some of which I explore here. The role of emotions needs to be brought into these investigations.

Emotion is typically conceived of as being irrational, or at the least non-rational. I contend, and culling support from other philosophical sources will argue, that rationality is really only fully intelligible, to the extent that it is fully intelligible, once we take the issue of emotions into consideration in our attempts to account for the phenomenon of rationality.

I have set out for myself, then, a broad two pronged problem involving emotions and rationality and their relationship to each other. I contend that if we hope to undertake to understand either one of these we need to do this in the light of the other. Hence I want to address myself here to why I think this is so. More specifically I want to focus on the essential role that emotions might play in normal (rational) cognition in creating salience and helping to limit the search for evidence.

Outline

In Chapter I, the INTRODUCTION, I have outlined the problem which I wish to address: the need to draw certain kinds of cognitive limits in order to be rational and the special role that emotions play in that process.

In Chapter II I outline a theory of rationality. I do this by considering views on rationality and epistemology and by trying to point out strengths and flaws in order to arrive at my own account of rationality. My own view emphasizes the problem of our limited capacity as "knowers" and how we can accept that and still pursue truth and knowledge. Central to this is the need to limit the search for evidence; we must accept our reasoning ability as limited and come to closure.

In Chapter III I consider some theories of emotion in order to help understand what role they might play in the concept of rationality I address in Chapter IV. There I consider how emotions play a specific essential role in rational cognition, giving special consideration to de Sousa's arguments for their essential role in creating salience, and in so doing helping us overcome both cognitive binds of indecision (Buridan's Ass) and the need to exclude possibly relevant evidence (the frame problem). Emotions contribute to limiting evidence on the level of functional cognitive architecture. I argue that while de Sousa is correct in a weak sense about the role of emotions in rational cognition, he attempts to give it a stronger sense than is defensible.

In APPENDIX I I address one of the questions that arises in Chapter II: What reasons do we have for holding reasons in esteem, or whether there is a rational basis for being rational.

In APPENDIX II, CONTEXT AND METHODOLOGY, I first consider two issues that I think are relevant to the context of this discussion: 1) the mind/body problem and 2) the free will/determinism debate. I call these "context" because I see them as having to do with how the problem I am addressing fits into a larger picture which includes the embodiment of mind. The free will/determinism debate comes up in trying to determine the nature of rationality. Earlier in Chapter IV, I argued against a conception of rationality as deterministic; a view which I think is dependent on an outmoded way of conceiving of rationality. Lastly, methodological lines, in the third section of APPENDIX II I consider different levels of description that we use in describing the functioning of minds, and how these different levels of description can be overlapped and used together to flesh out an account of mind, as is done in cognitive science.

II A THEORY OF RATIONALITY

1 Introduction

In this chapter I outline a theory of rationality through the use of various sources, appropriating or rejecting various aspects of their accounts of rationality. In the first three sections I try to distinguish my approach to rationality from what I think is a flawed, static approach to the subject before considering some contemporary accounts. I take issue with one of these accounts which suggests that defining rationality gets tied up with defining truth and that this has to do with a correspondence between language and reality. I maintain that understanding truth in this way will tend to lead us toward a justificational and static approach with all its ensuing errors. Truth is more adequately accounted for by a coherence theory. If truth is to have any meaning for us in our limited capacities it must be something other than correspondence - which is a more adequate position for a deity. Rationality demands of us that we try to increase our knowledge despite the fact that our evidence is inadequate. Our only way to assess its adequacy is to decide if it fits into our understanding of the world in a coherent way. A coherence approach to truth and the need to accept risk assessment as central to rationality (due to the predicament of reason) go hand in hand. Indeed, as Gilbert Harman argues, we must function with a coherence theory (at least in terms of our belief structures) and its inherent conservatism, because this is all our minds are capable of. Using evidence from cognitive psychology he argues that correspondence is simply beyond our cognitive capacities. Harman discusses principles of rationality which can help in the meta-cognitive assessment of our own reasoning, but we are still left in the predicament

of reason. Towards the end of the chapter I consider the limits to our meta-cognitive self-assessment. Before I leave off for Chapter III I point the way toward Chapter IV in which I consider how our functional architecture might take care of some of the cognitive problems with which we are faced. In other words, Chapter II discusses how we might try to look after certain cognitive problems. Chapter IV considers how they might look after themselves possibly with the help of emotion.

2 Naturalizing Rationality

In my discussion of the theory of rationality I will give consideration to a recent trend towards the naturalization of epistemology and theory of rationality. As I have suggested above, that which just happens to be the case when we look at how people reason and think will be of the utmost relevance to what we say about how people ought to think. This is so because what has been held true in ethics through the ages, namely that "ought" implies "can", must also hold true in epistemology. Insofar as epistemology is to form a normative guide for thought, and this should perhaps be more properly referred to as theory of rationality, it must give consideration to the ways people actually do think, and the ways in which they are capable of thinking. It has too often been the case that epistemology has suggested that people do things in their reasoning that individual human beings are not actually capable of, for reasons of limited memory capacity, limited time, inability to focus on a multiplicity of things at once, etc. There has also been a focus on the static manifested in a concern for the justification of beliefs. A naturalized epistemology, as in Harman and Code, focuses on and gives greater emphasis to the transitions in belief that we undertake or undergo. There is an implicit recognition of

the dynamic character of the mind of the human knower which leads to a greater recognition of, and emphasis on, the diachronic aspects of knowing as opposed to its synchronic aspects than epistemology had been wont to emphasize up to this point. How this naturalizing tendency influences our approach to the descriptions of rationality will become more apparent below.

3 A Note Regarding Irrationality

In trying to understand rationality I want to look at some philosophical notions of irrationality, not merely in order to begin to define rationality negatively, but also because the discussion I envision will help to define the territory of its positive discussion.

Bela Szabados discusses the nature of irrationality in his brief essay "A Note on Irrationality". He begins his discussion with a consideration of examples taken from other philosophers (Kai Nielsen and Kurt Baier) which they offer as paradigmatic examples of irrationality. Szabados quotes the following passage from Nielsen:

A man who for no reason at all, utterly ignored what he knew to be in his own interest would be behaving in a paradigmatically irrational manner. (Szabados,387)

The example taken from Baier is similar in that it is of an individual who undertakes self mutilation (cutting his own toes off) even though he knows that it would be in his own self interest not to do so. Szabados selects these examples which, though extreme, he does not see as being extraordinary in terms of the types of examples that philosophers are wont to offer as examples of irrationality. He wants to point up, however, that these are better seen as examples of madness than they are of irrationality. "They are beyond the reach of reason: their behaviour is

paradigmatic of insane behaviour." (Szabados,387) We have ways of accounting for madness or mad behaviour. We can, for instance, give causal explanations that function on a physical level, or perhaps functional explanations that function on a human level. These latter accounts might involve the notion of defense mechanisms. The point is the following:

...mad behaviour...hardly illustrates or serves as a paradigm of irrationality. Its advocacy is an instance of philosophical myopia. One is reminded here of Wittgenstein's perceptive diagnosis: "The main cause of philosophical disease - a one-sided diet of examples: one nourishes one's thinking with only one kind of example." [Wittgenstein,593](Szabados,388)

Irrationality, Szabados wants to say, has to do with the break down of rationality within a context of a normally rational individual. As Nicholas Rescher puts it: "One can only be actually irrational if one has the capacity for reason and proceeds to neglect, misuse, or abuse it."

(Rescher,198) Szabados draws the analogy of breaking rules while playing a game. One must first be playing the game first before one can be said to be breaking the rules. Furthermore, if an individual is to play the game then she must have some sense of the rules before the playing of the game will begin to be possible. Irrationality is breaking the rules of the game. Madness is not playing the game at all.

... the widespread and philosophically interesting cases of irrationality are those when a generally rational agent, with a concern for rationality, nay, in the name of rationality, believes or behaves irrationally: whence the need for self-deception and rationalisation. (Szabados,388)

Szabados sees as philosophically interesting the kinds of irrationality that will involve self-deception or rationalisation. The individual who realizes that his beliefs or actions are irrational and accepts them as such, according to Szabados, is already on his way to madness. At the other end of the scale is the individual who, upon recognizing the irrationality of

his beliefs or actions, abandons or reassesses them. We tend to call this type of irrationality a mistake. The important thing to note, however, is the distinction between what is irrational, and as such involves a context that is essentially rational, and what is madness, and as seems much more removed from a rational context.

I do not wish to address at this point the problem of whether the distinction between the irrational and the mad is one of kind or merely one of degree. Nonetheless, I think that Szabados' point about the distinguishing characteristic here of a rational context is worth noting. Certainly our attempts to account for mad behaviour are distinct from our attempts to account for a mere incident of irrationality.

We see then, along with Szabados, that madness is quite distinct from a mere incident of irrationality, and that the latter pre-supposes a context which is normally rational. Though the distinction between the irrational and the mad is not my key concern here, I hope the preceding discussion helps to define the territory for the discussion of rationality to follow, even if only negatively.

4 Defining Questions

Szabados draws an analogy between rationality and a game. The question, then, is, What are the rules of the game? What does or would a theory of rationality look like? This question leads to the question: "What questions does/should a theory of rationality ask/answer?" I will begin by considering what are reputed to be some of the most essential problems or questions for a theory of rationality as well as some of the central features of a typical conception of rationality. I will then go on to try to flesh out my initial bare bones account of the subject area.

Rationality is typically conceived of as a concept which has to do with the assessment of reasons. De Sousa says that rationality is a teleological notion and that only intentional acts or states are subject to rational assessment. (de Sousa,161) Any act or state that aims beyond itself or has directedness (or aboutness) is subject to rational assessment. This would include beliefs (about something), actions (toward some end, but not done for their own sake), judgments, desires (for something), and any other possible intentional act or state. We sometimes speak of persons being rational or not, but this, I maintain, is best understood as an ascription of a tendency to behave rationally with respect to the narrower parameters that I want to define here. It is a complication that yields little of value. Now the ways in which we might assess each of these things is quite distinct since there are quite different notions of rationality involved here. I presented above de Sousa's notion that rationality is a teleological concept. We can flesh this out when we consider what the proper end of each of these intentional states is. Thus, the proper end of belief is truth. Therefore a belief is rational insofar as

it achieves its proper end. But clearly this is too strong for we don't want to say that only those beliefs that we hold which turn out to be true are the ones that we hold rationally. For this reason we attempt to define some type of methodological considerations and say that those beliefs which are arrived at in a way that tends to lead to true beliefs are held rationally. The project becomes complicated when we try to define a method which tends to lead to true beliefs. Furthermore, it seems clear that our assessments of rationality with respect to beliefs will be quite distinct from our assessments of rationality with respect to actions. This distinction will become clearer in the following discussion of the views of de Sousa.

I suggested above that rationality had to do with the assessment of reasons. This leads to the meta-question concerning the justification of rationality: What reasons do we have for requiring reasons, or for holding reasons in esteem? There is a discussion that surrounds this question which I will address in APPENDIX I.

Another meta-question has to do with whether a theory of rationality should be descriptive or normative. Now while a purely descriptive theory might seem to make little sense since it is difficult to say what or who would be held up as the paradigmatic case, it also is the case that the way in which a normative theory is to be outlined needs to be tempered with some description of the ways in which people actually think and function. If we consider the widely accepted tenet in ethics that "ought implies can" there is little question as to why this should be the case. This has not always been accepted though, and philosophers have often outlined principles of rationality that give little consideration to the ways in which

people actually think. In more recent years, though, there has been a naturalizing tendency in accounts of rationality. Let me now try to flesh out an account of rationality that addresses these questions a little more fully, by considering a few contemporary accounts of the nature of rationality.

5 Cognitive versus Strategic Rationality: de Sousa

In *The Rationality of Emotion* Ronald de Sousa suggests that in trying to understand rationality we end up enmeshed in a triad of concepts that are inter-dependent for their meaning and intelligibility. These are: rationality, truth, and objectivity. Furthermore, objectivity is often understood insofar as it contrasts with subjectivity, which has several different meanings. (de Sousa,144) De Sousa writes:

Intuitively, truth has to do with a correspondence between language or thought and the real world, conceived of as existing independently of our conscious states. Objectivity seems more fundamental, because we might make sense of states other than thoughts or propositions that correspond to something in the real world. The perception of a certain quality, for example, can be objective, though we would not normally think of it as true. Truth, then, seems to be merely a species of objectivity. But objectivity must get part of its meaning from its contrast with subjectivity, which ... is not one thing but several. (de Sousa,144)

De Sousa goes on to suggest that the concept of rationality can be seen as derivative of these concepts of truth and objectivity insofar as it "... has to do with the strategies for the attainment of truth or objectivity." (de Sousa,144) What is meant by truth is, nonetheless, not universally agreed upon, and when uttered by a relativist this word may be understood quite independently of objectivity. De Sousa rejects relativism after considering Socrates' arguments against the kind of relativism espoused by Protagoras in the *Theaetetus* exemplified by the phrase "true for me". This kind of

position is often called subjectivism and is importantly different from a stronger and more defensible kind of relativism that maintains that truth is linked to communities rather than individuals. This brand of relativism lies somewhere in between subjectivism and the kind of correspondence theory of truth which de Sousa would seem to be espousing above.

Nonetheless de Sousa escapes from the possible problems of a correspondence position by taking a more semantic tack - defining truth in terms of objectivity, which he distinguishes from various kinds of subjectivity - and ultimately ends up focusing on subjective convergence. It is this notion of intersubjectivity which is required to maintain the possibility of meaningful linguistic communication.

After considerable efforts to outline the notions of truth, objectivity, and subjectivity which his theory of rationality presupposes, de Sousa outlines six "Principles of Rationality". These are as follows:

- (R1) Success. The formal object of a representational state defines that state's criterion of success, in terms of which the rationality of that state is assessed.
- (R2) Minimal Rationality. It is a necessary condition of an intentional state or event's being describable as categorially rational, that under some true description it can properly (though perhaps vacuously) be said to be evaluatively rational.
- (R3) Intentionality. The teleology implicit in rationality applies only to intentional acts or states.
- (R4) Origins. The assessment of rationality of any act or belief looks both forward to consequences, logical and causal, and backward to origins.
- (R5) Constraints. Rationality never prescribes, but only constrains, by proscribing inconsistency and distinctions without a difference.
- (R6) Cognitive and strategic rationality. A representational state can be assessed in terms of its probable *effects* (in the causal sense): this evaluates its *strategic* rationality, or utility. By contrast, a state is *cognitively* rational if it is arrived at in such a way as to be probably adequate to some actual state of the world that it purports to represent. (de Sousa, 336-7)

De Sousa sees rationality as a teleological concept that applies to

intentional states (R3). Any intentional state that is to be considered as rationally assessable, i.e., as categorially rational, must, under some description be rational, i.e., evaluatively rational (R2). Intentional states are assessed for their success insofar as they tend to achieve their proper formal object (R1). Thus, for example, the formal object of belief is truth, and a belief is evaluatively rational insofar as it tends to attain truth. Assessments of rationality should consider the likely causal effects of a state (R6) and thus should look forward to consequences (R6) (strategic rationality), and should look to origins (R4) to consider whether it is arrived at in such a way as to "be probably adequate to some actual state of the world" (R6) (cognitive rationality), i.e., likely to be true. De Sousa also wants to say that rationality leaves the domain of our inquiry open and only proscribes certain errors (R5). This principle is in conflict with Rescher's notion of a rationality that "speaks in didactic tones" (Rescher,9) as we will see below.

De Sousa here is distinguishing two different types of rationality. One he calls strategic, and the other he calls cognitive. This, I think, is done aptly. Consider that strategic rationality looks forward to consequences. For example, when I want to consider whether the act I am about to undertake is rational I will most certainly consider its likely consequences. It seems to me that I am less likely to look to origins when I am trying to consider the rationality of an act. Now, with regard to a belief it seems that the exact converse would be true. In this instance I would be more likely to assess the rationality of a belief in what de Sousa calls "cognitive" terms rather than strategic ones; I am more likely to consider the origins of the belief than its consequences in assessing its

rationality. Thus there is the makings of a division here between rational assessments for beliefs and judgments, and rational assessments for actions and wants, where beliefs and judgments have their own brand of rationality called cognitive rationality and actions and wants are assessed in terms of strategic rationality. This distinction, however, is not a thoroughgoing one.

Both standards of rationality can be defined in strategic terms. This explains the attractiveness of pragmatism. Truth, on this view, is just the goal or standard of success of the enterprise of belief. But we should not infer that the pragmatic is the more fundamental concept, because the tables can be turned. We could as easily think of successful action in terms of the epistemic goal, the concept of *matching*. (de Sousa, 163)

Now, while typically we can understand any given situation under one or the other description of rationality, there can be instances where the different descriptions of rationality urge us to quite distinct positions.

De Sousa uses the example of Pascal's wager and suggests that Pascal was arguing that we consider the rationality of holding a belief in god strategically rather than cognitively, namely that we should consider the consequences of this act of belief rather than the epistemic grounds for holding it. If we do this we discover that we should hold the belief even though we find that the epistemic reasons for holding it are quite weak. While this shows that indeed we do not always assess beliefs cognitively, I think that the belief in god is a peculiar instance in which a belief takes on the form of an overt act (quite distinct from the ordinary form of belief). This is so for the following reasons. If we are to believe in god because we are influenced by Pascal's wager we do this precisely so that god knows that we believe in his existence, since it is through god's knowledge of our belief that we receive the benefits (a strategic

consideration) of this belief. But, since god is an omniscient being, my holding of this belief is, for god, an overt act. I need do nothing more than hold the belief in order for god to know that I hold the belief. Typically, however, when we hold a belief it is not so much like an overt act and it does not have consequences in the same way an overt act does (though the belief in god in the instance of Pascal's wager does have the kinds of consequences of an overt act).

There is a further problem in the example of Pascal's wager in that it supposes that we can freely choose to either believe or disbelieve in the existence of god. That is, in order for us to believe in god for strategic reasons or out of a concern for consequences and not for cognitive or epistemic reasons it must be the case that belief is voluntary. The notion that belief is subject to the free choice of our will is at least controversial, if not unacceptable. But, it would seem that if belief is not voluntary then the notion of basing beliefs on strategic reasons is at least somewhat weakened, though there remains the position that believing is an act like any other (though not overt). Nonetheless, it seems that as the position of voluntary belief is eroded the distinction between cognitive and strategic rationality will be reinforced.

Pheroze S. Wadia in "Reasoning, Believing, and Willing" argues against the notion of voluntary belief. He calls the position he argues against the "voluntarist view of belief" by which he means, roughly, a position that maintains that it is "... within our power to believe, at least sometimes, contrary to our better judgment". (Wadia,233) What he means by this is something quite distinct from the voluntariness with which I undertake to investigate some given particular thing the truth of which I

wish to ascertain. Indeed, I undertake my investigations, and select the topics of such investigations voluntarily, and insofar as this is true there is a sense in which I choose what to believe. But, Wadia says,

... this is not at all the kind of position in regard to our freedom to believe that the proponents of what I have dubbed the Voluntarist Paradox have in mind. For the freedom to believe or not to believe a proposition that these philosophers have in mind is said to be operative even *after* the most thorough and honest assessment of the evidence in regard to its truth has been completed. The freedom it is claimed one has with respect to one's beliefs is said to be in a radical sense like the freedom one has with respect to one's actions. (Wadia,234)

Wadia goes on to suggest that "It would ... be natural to suppose that the most fundamental problem a critic should address in connection with such views is whether any belief is or ever can be brought about voluntarily" (Wadia,236) though he himself finds this approach to be unfruitful. What he does do is attempt a criticism of the notion of an ethics of belief whose possibility he thinks the voluntarist position presupposes. His argument is too involved to get into here but the sum of the issue is that the notion of the voluntariness of belief is at least controversial. Given convincing evidence for the truth of some position it is hard to conceive of how we can but help to accede to said position. There are, of course, a great range of evidential possibilities that are less than thoroughly convincing of the truth of some proposition and a whole range of doxastic attitudes that correspond to these evidential possibilities. The involvement of the will in determining belief states in these circumstances is likely much less problematic.

Now, the case of god is certainly one of those cases in which the evidence for the truth of the position is less than overwhelming, barring

occurrences of divine revelation. So does this mean that we can readily take up this belief by willing it? If we are in definite middle ground then it certainly seems that we might be able to do this, but if we believe that, on the contrary, there are good reasons to disavow this proposition, then it is hard to see how we can take up this position as a pure act of the will. Can I will to believe something that on evidential grounds I find untenable for the reason that I wish to hold the belief for strategic reasons - because I wish to have the benefits of the holding of this belief? De Sousa sums up the situation as follows:

Is there a criterion, in cases like Pascal's bet, for what is rational *absolutely*? I see no prospect of finding one. The question can always be rephrased in each of two ways, predetermining opposite answers. I can ask, "What should I do: believe it, or not?" - which calls for the answer, "If it's useful to do so, then believe it." Or I can ask, "What should I take it to be: true or false?" - which demands the answer, "If it's highly improbable, then don't believe it." (de Sousa, 164)

De Sousa sees any possible solution to the situation as involving an intractable dichotomy. Insofar as I conceive of belief as an action like other actions, I consider its rationality strategically and, insofar as I consider belief as an intentional state with truth as its proper object then I consider its rationality cognitively.

I have tried, above, to give reasons why we should not consider belief to be like other actions and why, furthermore, it may not be acceptable to suppose that we can volitionally take up beliefs for which we lack an evidential basis. De Sousa's intuitions may, nonetheless, be correct for situations in which evidence is not compelling one way or the other. It seems to me, however, that there must be a point at which it becomes unacceptable to override epistemological or evidential considerations for

strategic considerations when determining what to believe. Surely there comes an evidential point at which we simply will (or will not) believe some proposition no matter how useful a belief in said proposition might be (or at least there is a point at which this is the rational thing to do). For example, there are many reasons why it is useful to believe that my wife was not unfaithful to me when she took that business trip last year: this belief gives me peace of mind; it allows me to remain open and trusting of her; even if she was unfaithful to me then I know that it would have been an isolated incident and it would not change my feelings for her in the long run; in general it serves to facilitate my own happiness to believe that she was not unfaithful to me then. Yet when the evidence mounts that she did indeed have an affair during her business trip I accede to this evidence even though in many respects it may be more useful to me to continue to believe the opposite. In the face of mounting evidence I cannot simply choose to believe what it is more useful to believe. On the other hand, if for some vague reasons I have suspicions that she may have had an affair at that time (she seemed strangely distant for a couple of days upon her return), yet I lack any real substantive evidence, then I may choose to believe what it is more useful to believe, that is, that she did not have an affair. I do this for the above-mentioned strategic reasons. Furthermore, I likely come up with an account for myself as to why this is the cognitively rational thing to do. (She was distant at that time because she was pre-occupied with her work.)

William James in "The Will to Believe" arrives at a very similar position with respect to the possibility of volitional belief. He suggests that when we lack the evidence to determine the truth of something our

"willing nature" may determine our belief.

When I say "willing nature", I do not mean only such deliberate volitions as may have set up habits of belief that we may not now escape from - I mean all such factors of belief as fear and hope, prejudice and passion, imitation and partisanship, the circumpressure of our caste and set. As a matter of fact we find ourselves believing, we hardly know how or why. (James 1980a,191-192)

Clearly, James' notion of volitional belief here involves a psychology of the influence of underlying mental states in determining what belief states we will, in fact, come to hold. This is not to say that he dismisses the notion of volitional belief in the more straightforward sense of consciously choosing to believe something, though we certainly cannot choose to believe things that are "contrary to our better judgment" (i.e. James rejects what Wadia calls the voluntarist position.) Only in cases where we lack the evidential grounds may we choose to believe one thing over the other.

James says,

Wherever the option between losing truth and gaining it is not momentous, we can throw the chance of gaining truth away, and at any rate save ourselves from any chance of believing falsehood, by not making up our minds at all till objective evidence has come. In scientific questions, this is almost always the case; and even in human affairs in general, the need of acting is seldom so urgent that a false belief to act on is better than no belief at all. (James 1980a,199-200)

The other side of the coin, however, is that whenever we are faced with a decision to make about an option that is "momentous" and not merely trivial, that is forced, i.e., not avoidable, and where the options are living options, i.e., the options have "some appeal, however small, to your belief" (James 1980a,187), then we must decide on some other basis than intellectual grounds when these are not available. I will delve further into James' view of the nature of belief (and truth) below.

The point of the above discussion of Pascal's wager was to attempt

to argue that we should perhaps try to reinforce the distinction between the two different types of rationality that de Sousa describes, namely cognitive and strategic rationality.

6 Practical and Evaluative Rationality: Rescher

Nicholas Rescher in his book *Rationality* seems amenable to the drawing of this kind of distinction and indeed wants to make the grain even finer and divide strategic rationality into two different, more specific, types. He suggests the following break down:

1. COGNITIVE RATIONALITY: What to believe or accept? What to maintain regarding states of affairs in both the formal and the empirical domains. Product: factual contentions (beliefs).
2. PRACTICAL RATIONALITY: What to do or perform? What to decide regarding actions. Product: action recommendations (injunctions).
3. EVALUATIVE RATIONALITY: What to prefer or prize? What to adopt regarding goals or ends. Product: evaluations (appraisals). (Rescher,3)

Here Rescher avoids the suggestion (which de Sousa makes) that a given type of intentional state might be subject to different types of rational assessment and suggests rather that each of these different types of intentional state has its own type of rationality. He does seem to be, nonetheless, more in accord than in disagreement with de Sousa about the nature of rationality. He thinks that rationality is best understood as a means/end notion. Thus he says:

Rationality thus calls for appropriate resolutions intelligently arrived at and sensibly implemented. It is geared to the sensible pursuit of appropriate ends. (Here 'ends' is used as a generic term covering both concrete goals and more diffuse values.) Rationality is thus a two-sided Janus-faced conception. On the side of means, it reflects a pragmatic concern for efficient process, while on the side of the 'appropriateness of ends' it reflects a value-geared concern for the product. (Moreover, the acceptability of the means themselves also enters in.) (Rescher,6)

Rationality has to do with the pursuit of some end which is rationally selected, by rationally selecting means which are "sensibly implemented". Rescher here suggests that "efficient process" is of primary concern when it comes to assessing means. Further on he writes:

The logician takes consistency to be the bulwark of rationality, the scientist evidential cogency, the economist efficiency. All are right, but each only partially so. Each focuses upon what is no more than a part of reason - that one particular part of intelligent procedure that is of primary importance for his own field. The crucial point is that rationality as such is something complex and many-sided - though all of its parts are embraced within the one, overall generic formula that rationality consists in the intelligent conduct of one's affairs. (Rescher,8-9)

What any given individual will take to be most essential to rationality will depend upon her interests, yet, Rescher wants to say, there is something that ties these notions of rationality together and it is precisely that "rationality consists in the intelligent conduct of one's affairs." Rescher's notion of rationality is a practical one, having to do with "conduct". This conception of the nature of rationality could not be much broader. Rationality here is essentially "intelligence" - put to purpose in the conduct of one's affairs. "Rationality is not *just* a matter of thought, but of action as well." (Rescher,3)

Rescher sees his theory of rationality as being distinct from the standard fare in rationality theory (especially as it is rooted in specific disciplines such as economics, logic or decision theory) in two respects: 1) Its emphasis on the breadth of rationality, and 2) Its claim of the inherently normative aspect to rationality:

Every specialty seems to opt for some narrow desideratum as the definitive feature of reason. In fact, however, rationality is something far-reaching and much-inclusive, and not merely a particular delimited good to be achieved by narrow technical

means ... rationality is as wide-ranging and complex as the domain of intelligence at large. The central thesis of the book is that only a normative theory of rationality can be adequate to the complexities of the subject. (Rescher,vii-viii)

Rescher emphasizes the breadth of the notion of rationality and warns against an "... overly cerebral conception [of rationality which] represents a far too narrow and one-sided view of reason's domain. Rationality is broad and comprehensive." (Rescher,9) To this end Rescher notes that rationality involves a breadth of context. Thus, for instance, it is sometimes in the interest of rationality to act in an irrational way in the short term. This may, for example, serve a broader tactical concern such as confusing the enemy, or enticing a lover by playing hard to get. Furthermore, Rescher suggests, "Feelings are not generally a matter of reasoning, but they are certainly not outside the province of reason." (Rescher,9)

7 The Predicament of Reason (The Injunction to Believe and Risk Assessment)

What Rescher calls the "predicament of reason" pertains to the breadth of context as well as the notion of rationality as a practical notion. The predicament arises because as rational beings we must function not with ideal circumstances (full information, etc.), but only with what we have available. We base our decisions upon some kind of assessment as to the adequacy of our information. Reason's predicament is that it is virtually always the case that there is some possible piece of additional information that will render our decisions (determinations of truth, of a course of action, of what is valuable, etc.) virtually contrary to what rationality calls for.

We, nonetheless, must persist in going about our reasoning as we do,

for the only alternative, Rescher suggests, is scepticism. The sceptic, he suggests, makes the error of inappropriately avoiding risk. While it is true that when acting rationally there is always the risk that we will err, it is an even greater error to avoid reasoning at all simply because of the possibility of error. "Rationality is realistic; it does not require more than is possible." (Rescher,7), but,

Rationality makes demands upon us. It speaks in didactic tones: this or that is what you *should* do. Its declarations have a normative force, enjoining us as to how we ought to go about settling questions of what to believe, do, or value. 'These are the strongest reasons for belief, action, and evaluation - be governed by them!' (Rescher,9-10)

Rescher is in clear agreement with the suggestion that I made above that rationality, while being a normative concept needs to be normative in a way that is conscious of the limitations of the human animal. What then, for Rescher, does rationality demand of us? He suggests, in the passage above, that rationality demands that we be governed by the strongest reasons. This remains a very broad understanding of rationality and needs to be fleshed out a little further.

Consider the example of the sceptic. Rescher suggests that the sceptic has fallen prey to unreasonable risk avoidance. The sceptic would want to maintain that his position is rational since the aim of belief is truth, and with respect to this he never errs. De Sousa might be amenable to this position. Recall his principle of rationality (R5) that states that rationality is not prescriptive but only proscribes "inconsistency and distinctions without a difference." (de Sousa,163) If rationality is not prescriptive it cannot demand of us that we avoid scepticism and seek knowledge or information. This, however, is not

Rescher's position, and with good reason. He writes:

Reason's commitment to the cognitive enterprise of inquiry is absolute and establishes an insatiable demand for extending and deepening the range of our information. Reason cannot leave well enough alone, but insists upon a continual enhancement in the range and depth of our understanding of ourselves and of the world about us. By rejecting the very possibility of securing trustworthy information in factual matters, scepticism sets up a purportedly decisive obstacle to implementing these aims of reason. (Rescher,48)

For Rescher, unlike for de Sousa, it is not enough to merely avoid inconsistency and false distinctions. Reason provides a positive impetus that entreats us to broaden and deepen our knowledge. This brings the issue of risk to the fore. Rescher has suggested that the sceptic has been stultified in his search for knowledge by unreasonable risk avoidance. In fact the issue is more complicated. Rescher describes two types of misfortunes that one might want to avoid:

Misfortunes of the 1st kind: We reject something that, as it turns out, we should have accepted. We decline to 'take the chance' and avoid running the risk at issue, but things turn out favourably after all, and we 'lose out on the gamble'.
 Misfortunes of the 2nd kind: We accept something that, as it turns out, we should have rejected. We do 'take the chance' and run the risk at issue, but things go wrong, and we 'lose the gamble'. (Rescher,57-8)

Now while the sceptic does particularly well at avoiding misfortunes of the 2nd kind (at least cognitively speaking) he does utterly horrifically at avoiding misfortunes of the 1st kind. There is, of course, also a correlate of the sceptic with respect to reasoning about what to do and value. Conversely there are risk seekers (in the various realms of reasoning) who are in the opposite situation to that of the sceptic, viz., they do not frequently suffer from lost opportunities (misfortunes of the 2nd kind) and they do frequently suffer negative outcomes. On the scale that leads from the risk avoider to the risk seeker are those individuals who assess

individual situations and calculate risks. The scale moves from the more cautious to the more daring calculators. Rescher suggests that either extreme, risk avoiding or risk seeking is not rationally optimal.

The rationally sensible thing to do is clearly to adopt a policy that minimizes misfortunes overall. ... The sensible thing is to adopt the 'middle-of-the-road' policy of risk calculation, acting as best we can to balance the positive risks of outright loss against the negative ones of lost opportunity. In the cognitive case ... the sceptic succeeds splendidly in averting misfortunes of the second kind. By accepting nothing, he accepts nothing false. But, of course, he loses out on the opportunity to obtain any sort of information. ...[and]... After all, what we want in inquiry - the object of the whole enterprise - is information. (Rescher,58-59)

This is much in keeping in line with Rescher's pragmatic bent to an analysis of rationality. He dismisses the possibility of avoiding risk outright and suggests that the attempt to do so only leads to the misfortune of lost opportunity. Thus, since risk is inherent in our enterprises, whether it is in what we do, believe or value, he suggests that we take a route that minimizes misfortunes of both types. To this end we follow a middle track. We proceed with the task at hand, but we do so in a way that attempts both to maximize benefits and minimize unnecessary risks.

8 Evaluative Rationality

One of the key things to note about this purview of rationality is that it is concerned not merely with how we should achieve certain ends (how to achieve the proper end of belief, for instance, i.e., truth), but with which ends should be sought. Rationality is not a value neutral notion that merely lays over the surface. It is not a road map that tells us how to get to wherever we might wish to go. It comes along for the ride and when the rest stop takes too long it urges us back onto the road lest we

never get anywhere.

But not only does rationality urge us onward when we fail to proceed along the road, it also chides us for heading toward the wrong destination. This is what Rescher refers to as evaluative rationality. Now, clearly in the case of certain pre-established ends what we will be primarily interested in is the means to those ends. Thus, for instance, true beliefs are ends, the value of which we are not going to question, and are the subject of cognitive rationality. Nonetheless there are other cases of deliberative reasoning that do not merely pertain to means but to the ends themselves. Taking off from an example from Aristotle, Rescher writes:

... the doctor does not deliberate about treating illness - that choice is already settled, included as part of his decision to become a doctor. But, a young person may well deliberate about whether to become a doctor in the first place - reflecting on whether this would be something good for him, given his abilities, skills, interests and so on. And this sort of deliberation is not a question of means to pre-established ends at all. The long and the short of it is that there are two very different sorts of deliberations: cognitive deliberations regarding matters of information and evaluative deliberations regarding matters of value. (Rescher, 92-93)

This notion of evaluative rationality is in strong contradistinction to a strain of thought in the West that insists on the instrumentality of reason and is epitomized perhaps by Hume in his statement:

It is not contrary to reason to prefer the destruction of the whole world to the scratching of my finger. It is not contrary to reason for me to choose my total ruin. ... It is as little contrary to reason to prefer even my own acknowledged lesser good to my greater, and to have a more ardent affection for the former than the latter. (Hume, *Treatise*, bk II, pt iii, sect 3)

Rescher, to the contrary, suggests that it is well within the domain of reason to analyze and evaluate preferences and it offers us not only

determinations of what we should believe but also of what we should value.

... the genuinely rational person is the one who proceeds in situations of choice by asking himself not the introspective question 'What do I prefer?' but the objective question 'What is to be deemed preferable? What ought I to prefer?' Rational comportment does not just call for desire satisfaction, it demands desire management as well. The question of appropriateness is crucial. ... For rationality, the crucial question is that of the true value of the item at issue. ... The normative aspect is ineliminable. There is an indissoluble connection between the true value of something (its being good or right or useful) and its being rational to choose or prefer this thing. Evaluation is at the very heart and core of rationality. For, rationality is a matter of balancing costs and benefits - of best serving our overall interests. (Rescher, 99-100)

Costs and benefits, the practical element, the notion of the inherent risk of some type of misfortune, all are involved in Rescher's understanding of evaluative rationality. But in order to appreciate the real costs and benefits we must distinguish, as Aristotle does, between what is really good and those things that are only apparently good. As Aristotle notes, this is tied up with human nature. There are certain types of goods that are goods for human beings and thus when I desire something I can evaluate the appropriateness of my desire, with a notion of human nature in mind.

9 Rationality and Principles

Rescher goes on to suggest that with an understanding of human nature we can arrive at universal principles which will legitimate specific interests insofar as they fall within the domain of the covering principle.

But just what is it that is in a person's real (best) interests? Partly, this is indeed a matter of meeting the needs that people universally have in common - health, normal functioning of body and mind, adequate resources, human companionship and affection, and so on. Partly it is a matter of the particular role one plays: co-operative children are in the interests of a parent ... Partly, it is a matter of what one simply happens to want. ... But these want related interests are valid only by virtue of their relation to universal interests. ... Any valid specific interest must fall within the

scope of an appropriate, universal covering principle of interest legitimation. ... The fact that X wants A remains a mere motive for his action in pursuing A (in contradistinction to a reason) until such time as it is rationalized through the fact that X recognizes A to have the desirable feature F, which is not just something that X wants, but is something that any and every (reasonable) person would want [in relevantly similar circumstances]. (Rescher,100-101)

Harvey Siegel places a similar, if not greater, emphasis on the notion of universal principles in his understanding of the nature of rationality. To be rational, for Siegel, is to be appropriately moved by (appropriate) reasons for our beliefs and actions. This requires, however, that we establish some criterion of appropriateness since any old reason will not do. Reasons require principles for their legitimation. "[T]o seek reasons ... is to recognize and commit oneself to *principles*, for, as R.S. Peters puts it, 'principles are needed to determine the relevance [and strength] of reasons'." (Siegel,33) Siegel goes on to emphasize the importance of principles for the assessment of reasons. Siegel considers an example of a school boy kept after class by the teacher because he was disruptive in class. The student's disruptiveness is the *reason* for keeping him after class. He writes:

For this properly to count as a reason, the teacher must be committed to some *principle* which licenses or backs that reason, i.e. establishes it as a bona fide reason, e.g. "All disruptive behaviour warrants keeping students after class" ... which must be consistently applied to cases [which are relevantly similar]. If the teacher is not committed to some such principle, then her putative reason for keeping Johnny after class does not constitute a genuine reason. (Siegel,33)

Siegel's view of the relevance of principles to the legitimacy of reasons seems correct. If the teacher does not act in this way as a matter of policy or principle (with some sense of consistency and regularity) then we may well indeed wish to question whether that was the genuine reason

for the action. For example, if there are four students consistently disturbing the class but only Johnny is detained, then we would need to look for some further reason why Johnny is detained and not the others; perhaps the teacher has some reason to believe that Johnny will benefit from this type of discipline whereas the others will not; perhaps she is acting on some principle that she is not justified to act upon (and thus is acting irrationally) - she detains students she doesn't like, even though there is no basis for this dislike. It is almost impossible to imagine that there would be simply no reason underlying Johnny's detention for this would seem to transcend the realm of irrationality and approach that of lunacy. Thus what truly needs to be assessed with regard to rationality is the relevance (appropriateness) of reasons, and the universalizability of the principle which purports to legitimate the reason. The principle must be valid; validity of this sort, as I have already noted, gets cashed out in terms of human nature and genuine human interests.

10 The Interdependence of Different Types of Reason

Reasons, then, must be rooted in principles in order to be legitimated. This assures that not just any reason counts as a good reason. But, furthermore, the different types of reasoning each make reference to the other and thus, though we can divide reasoning along the lines determined by the particular end of the given piece of reasoning, there is, nonetheless, a systemic unity to reason. For instance, in cognitive reasoning (reasoning about what to believe) both practical reasoning and evaluative reasoning will play a role. Rescher suggests that this is true anytime we arrive at new information from experience, which he calls the categorical aspect of cognitive rationality.

Cognitive reason as such has two sides, the hypothetical and the categorical. The hypothetical proceeds by way of: 'If you accept P, then you must (or alternatively, cannot) in due consistency also accept Q.' Its conditionalized, if-then aspect does not involve us in any substantive commitments whatever. But of course it also provides no substantive information. Only in the categorically committed use of cognitive reason do we obtain outright information ... this categorical dimension of cognitive reason has an ineliminably practical basis. ... To accept a thesis is to do something - even mental action is itself a kind of action, and belief-formation a kind of praxis. (Rescher,123-124)

He goes on to describe the dependence of cognitive reason on evaluative reason. He writes:

Once one has certain facts at one's disposal, one can, of course, proceed to derive others from them by logical inference. But, one must make a start at the inferential venture by way of accepting some facts not on the basis of other facts, but on the basis of mere indications. We cannot proceed in factual matters without making some non-discursive judgments of fact, such 'acts of acceptance' must be guided by evaluative processes. The stance that 'The indications are indeed strong enough to warrant acceptance in the particular case at hand' presupposes judgments along the lines of 'This issue is important enough for us to chance a resolution on present evidence notwithstanding the inherent risk of error'. (Rescher,124-125)

We must determine the value of some piece of information in order to determine what level of risk (of error) we should be willing to take in order to obtain the desired information. Rescher goes on to describe in some detail the other interdependencies that make up the systemic unity of reason, but I won't follow those arguments here.

Rescher sees the different types of rationality functioning in different ways but also being interdependent and forming a systemic whole. This would seem to be distinct from de Sousa's conception of the different types of rationality as being essentially different strategies with which we might approach the same problems depending on the question that we are trying to answer. But, how, for de Sousa, do we decide which question we

should be answering? How do we decide whether we should be asking, "What should I do?" or "What should I believe?" How do we decide whether we should be approaching a particular problem presented to our reason strategically or cognitively. It would seem that this is ultimately an evaluative question, and in this respect Rescher's account of rationality can address itself to that evaluative question much more readily than de Sousa's account of rationality.

There is much in Rescher's account of rationality that is similar to what William James had to say in 1896 in an essay entitled "The Will to Believe" which I alluded to briefly above. I will now consider James' view of reason, belief and truth in that essay.

11 James' View of Reason, Belief and Truth

The postulate that there is truth, and that it is the destiny of our minds to attain it, we are deliberately resolving to make, though the sceptic will not make it. We part company with him, therefore, absolutely, at this point. But the faith that truth exists, and that our minds can find it, may be held in two ways. We may talk of the empiricist way and the absolutist way of believing in truth. The absolutists in this matter say that we not only can attain to knowing truth, but we can know when we have attained to knowing it; while the empiricists think that although we may attain it, we cannot infallibly know when. To know is one thing, and to know for certain that we know is another. (James 1980a,194)

James asserts that we can essentially will to continue on in our search for truth. Yet, he suggests, to return to my analogy of journeying down a road, that we should continue down the road hoping that it is the right one, but never really being sure that it is. All we can be sure of is that the road we were on a way back was the right road, though we weren't sure of it at the time. It is a journey through an undefined territory. In the journey of investigation in search of truth we are wandering through the Australian outback, or perhaps better, a giant unmarked, unlit parking

lot. There is a multiplicity of possible tracks to follow but whether we are actually following a better one only comes to light when we see at a later point where it has led us. And I say "a better one" and not "the right one" because it is difficult to say what the latter would mean in this context. James' notion of truth does not seem to be of something with a clearly fixed nature.

I live, to be sure, by the practical faith that we must go on experiencing and thinking over our experience, for only thus can our opinions grow *more true*; but to hold any one of them - I absolutely do not care which - as if it could never be reinterpretable or corrigible, I believe to be a tremendously mistaken attitude. (James 1980a, 196, my italics)

It is clear, then, that for James truth is not absolute. Things can be more true or less true. He intends to contrast his view with that of (what he calls) the absolutists. With respect to truth the absolutists would seem to include everybody else in the philosophical community (and likely encompasses our common sense notion of truth as well). We typically conceive of truth in an absolute way. We say that a proposition is either true or false, not more or less true. For James, though, it is precisely in this sacrifice of our absolutist notion of truth that we really open up the possibility of attaining truth. When we are willing to approach truth gradually, getting closer little by little, rather than insisting on the perfection of truth from the outset, it allows us to pursue our investigations and not be stifled.

... when as empiricists we give up the doctrine of objective certitude, we do not thereby give up the quest or hope of truth itself. We still pin our faith on its existence, and still believe that we gain an ever better position toward it by systematically rolling up experiences and think. Our great difference from the scholastic lies in the way we face. The strength of his system lies in the principles, the origin ... for us the strength is in the outcome, the upshot ... It matters not

to an empiricist from what quarter an hypothesis may come to him ... if the total drift of thinking tends to confirm it that is what he means by it being true. (James 1980a,197-8)

James has realistic expectations for our investigations. They should aspire to the truth, not strike it with the first blow with one fell swoop. In order to do this, though, we must avoid the pitfall of the sceptic in curtailing investigation too soon.

Believe truth! Shun error! - these, we see, are two materially different laws; and by choosing between them we may end by coloring differently our whole intellectual life. We may regard the chase for truth as paramount and the avoidance of error as secondary; or we may, on the other hand, treat the avoidance of error as more imperative, and let truth take its chance. Clifford ... exhorts us to the latter course. Believe nothing, he tells us, keep your mind in suspense forever, rather than by closing it on insufficient evidence incur the awful risk of believing lies. You, on the other hand, may think that the risk of being in error is a very small matter when compared with the blessings of real knowledge, and be ready to be duped many times in your investigation rather than postpone indefinitely the chance of guessing true. ... Our errors are surely not such awfully solemn things. In a world where we are so certain to incur them in spite of all our caution, a certain lightness of heart seems healthier than this excessive nervousness on their behalf. (James 1980a,198-99)

James notes here the possibility of two radically distinct approaches to "our duty in the matter of opinion". (James 1980a,198) His discussion would seem to presage Rescher's, above, where he refers to a similar kind of distinction. James here offers a stronger emphasis than does Rescher on the inherent risk involved in the pursuit of truth. James is much in agreement with Rescher, though, with respect to their analyses of the problem of scepticism. James says:

Scepticism, then, is not avoidance of option; it is option of a certain particular kind of risk. Better risk loss of truth than chance of error ... (James 1980a,205)

He also would seem to anticipate Rescher's idea of the "predicament of reason" when he notes that one of Clifford's key problems with coming

to opinion is that we run the risk of closing on insufficient evidence, hence making an error. Rescher would note that when we reach closure there is always a risk that some additional information in the future will render our opinion erroneous. According to Rescher our opinion would be rendered erroneous. It would seem, however, that for James, at least under certain circumstances, when a certain opinion is superseded that does not mean that the superseded opinion is rendered false, but merely that it has been replaced by something that is more true. It lacks the absolutist's judgment of error from the eternal perspective. The "empiricist" (pragmatist), as James calls him, accepts that truth is something that is gradually moved towards and thus when an opinion is superseded it does not face the absolutist's condemnation for having been an error. It was, rather, a useful truth that has served its purpose because it aided us to come to a position that is more true.

Now, James' notion that we must continue on the path of investigation is a moderate view.

Wherever the option between losing truth and gaining it is not momentous, we can throw the chance of gaining truth away, and at any rate save ourselves from any chance of believing falsehood, by not making up our minds at all till objective evidence has come. In scientific questions, this is almost always the case; and even in human affairs in general, the need of acting is seldom so urgent that a false belief to act on is better than no belief at all. ... wherever there is no forced option, the dispassionately judicial intellect with no pet hypothesis, saving us, as it does, from dupery at any rate, ought to be our ideal. (James 1980a,199-200)

We do not usually forge ahead lacking any good evidence when we can simply wait for further evidence. James suggests that when we are faced with certain kinds of options, viz., living, forced, momentous options (to put it briefly: things that are of the utmost importance to us, that we

cannot delay deciding upon), then we may decide based upon something other than an evidential basis. He calls this our *passional* nature. He says:

Our *passional* nature not only lawfully may, but must, decide an option between propositions, whenever it is a genuine option that cannot by its nature be decided on intellectual grounds; for to say, under such circumstances, "Do not decide, but leave the question open," is itself a *passional* decision - just like deciding yes or no - and is attended with the same risk of losing truth. (James 1980a,193)

So, the strength of evidence that we will require to make a decision about whether to believe some proposition or not will have everything to do with our personal need to decide upon the issue. We cannot simply define rules of evidence and suggest that whenever these rules are satisfied we can decide and that when they are not we should suspend belief. The impact of the opinion upon us will be a key determining factor as to whether we can suspend belief in this way.

In a later essay, "What Pragmatism Means" (1907), James clarifies his conception of truth and in so doing ties in methodological/ psychological considerations in discussing the nature of belief revision.

He begins by outlining the importance of considering consequences or positive results according to the pragmatic method.

The pragmatic method ... tr[ies] to interpret each notion by tracing its respective practical consequences. What difference would it practically make to anyone if this notion rather than that one were true? If no practical difference whatever can be traced, then the alternatives mean practically the same thing, and all dispute is idle. (James 1980b,210)

And further on he adds pointedly:

There can be no difference anywhere that doesn't make a difference elsewhere - no difference in abstract truth that doesn't express itself in a difference in concrete fact and in conduct consequent upon that fact, imposed on somebody, somehow, somewhere, and somewhen. (James 1980b,212)

In some way, then, our abstract notions must be effective in our lives. Our ideas or notions, abstract truths, will ultimately prove to be things that we use - tools of a sort. Indeed, in certain respects an instrumentalist view is inherent in pragmatism, but this is not to say that a pragmatist is even going to be willing to enter into the instrumentalism/realism debate. Whether our abstractions correspond to some "real" object is a debate that can only be engaged in if we presuppose the kind of absolutist notion of truth to which I have referred. The pragmatist is beyond this debate, being an instrumentalist in the sense that she insists that the measure of a concept's worth (truth) is in its usefulness. Thus James argues that:

... an idea is "true" so long as to believe it is profitable to our lives. That it is *good*, for as much as it profits, you will gladly admit. If what we do by its aid is good, you will allow the idea itself to be good in so far forth, for we are the better for possessing it. But is it not a strange misuse of the word "truth", you will say, to call ideas also "true" for this reason? ... Let me now say only this, that truth is only one species of good, and not, as is usually supposed, a category distinct from good, and coordinate with it. *The true is the name of whatever proves itself to be good in the way of belief, and good, too, for definite, assignable reasons.* Surely you must admit this, that if there were no good for life in true ideas ... then the current notion that truth is divine and precious, and its pursuit a duty could never have grown up ... (James 1980b,223)

In suggesting that "truth is only one species of good" it looks like James might get himself into a tight jam with the example I presented above - my trying to determine what to believe when my wife comes home from a business trip and seems distant. In discussing that example I suggested that lacking good reasons to believe she had had an affair I can choose to believe, for what de Sousa calls strategic reasons, that she did not. I am believing that she did not have an affair because in some sense it is

better, i.e., "profitable" for my life, for me to believe it. Does this mean, then, that my belief is true in James' notion of truth, even if it turns out to be the case that my wife did in fact have an affair? I don't think that James' notion of being profitable falls victim to this kind of example because, even though it is true that James wants to avoid the so-called absolutist notions of truth and a strict correspondence theory of truth, this conception of truth really has to do with more theoretical notions.

In clarifying his notion of pragmatism by contrasting it with empiricism James takes note of pragmatism's commitment to facts.

Now pragmatism, devoted though she be to facts, has no such materialistic bias as ordinary empiricism labors under. Moreover, she has no objection whatever to the realizing of abstractions, so long as you get about among particulars with their aid and they actually carry you somewhere. (James 1980b,222)

The notion of facts is an important one. Much of what we know has to do with mere facts. "Is this a dagger I see before me?" Well, it's got a blade and a handle, it has the general appearance of a fighting weapon and not a kitchen tool and is not as long as a sword. Yes, it's a dagger. Period. We go into the courts and they ask for "the facts". If my wife had sex with another man then that is a fact. What exactly her motivation was, though, is not likely a factual issue. Human motivations can be quite complicated and very often are not even known to those who are moved. We are not completely transparent to ourselves and consequently human motivations are an area of theoretical investigation, not matters of fact. I explain my wife's behaviour by saying that she was trying, though not consciously, to jolt me out of my indifference to her because of my preoccupation with my work. Is this a useful explanation? If I pay her more heed and she continues to have an affair then this is not likely a

useful account of my wife's behaviour, or it at least needs to be expanded upon.

The pragmatist clings to facts and concreteness, observes truth at its work in particular cases, and generalizes. Truth, for him, becomes a class name for all sorts of definite working values in experience. (James 1980b,220)

The notion of truth, like all my other abstract notions, must be useful. It helps me get around in the world. Getting around in the world cognitively (not that we can really separate this from action) involves forming beliefs. James has a certain view about the acquisition of beliefs and the modification of the fabric of beliefs.

12 Conservatism, Coherence and the Fabric of Belief

James suggests that one comes to form "new opinions" when one's old opinions, upon meeting a new experience, are strained. Something comes along with which one's stock of opinions does not mesh.

Somebody contradicts them; or in a reflective moment he discovers that they contradict each other; or he hears of facts with which they are incompatible; or desires arise in him which they cease to satisfy. The result is an inward trouble ... from which he seeks to escape by modifying his previous mass of opinions. He saves as much of it as he can, for in this matter of belief we are all extreme conservatives. So he tries to change first this opinion then that (for they resist change very variously), until at last some new idea comes up which he can graft upon the ancient stock with a minimum of disturbance of the latter ... This new idea is then adopted as the true one. It preserves the older stock of truths with a minimum of modification, stretching them just enough to make them admit the novelty ... (James 1980b,216-217)

This notion of a stock of beliefs, a fabric as it were, that we stretch a bit and modify in order to squeeze in new truths is a powerful notion that comes to hold a great deal of importance in our understanding of human cognition. Thus W.V.O. Quine in 1951 writes the following:

The totality of our so called knowledge or beliefs, from the most casual matters of geography and history to the

profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges. Or, to change the figure, total science is like a field of force whose boundary conditions are experience. A conflict with experience at the periphery occasions readjustments in the interior of the field. Truth values have to be redistributed over some of our statements. Reevaluation of some statements entails reevaluation of others, because of their logical interconnections - the logical laws being in turn simply certain further statements of the system, certain further elements of the field. Having reevaluated one statement we must reevaluate some others, which may be the statements of logical connections themselves. But the total field is so underdetermined by its boundary conditions, experience, that there is much latitude of choice as to what statements to reevaluate in the light of any single contrary experience. No particular experiences are linked with any particular statements in the interior of the field, except indirectly through considerations of equilibrium affecting the field as a whole. (Quine, 42-43)

Quine here is arguing against the notion that truth value resides in individual statements or propositions and that what we do to determine the truth of a statement is simply compare it to our experience. It is only as a unified whole that our beliefs face the test of experience and he suggests that this is so in an "underdetermined" way which leaves a great deal of latitude as to what we might reevaluate given a conflict with experience. It is not simply going to be one statement, because even this will force us to stretch and readjust the fabric of beliefs, which may be done in many different possible ways. What's more, we may make adjustments in what Quine refers to as "the statements of logical connections themselves". This may mean that the weave of the fabric is somewhat altered (there is a change in the connections, simpliciter), or it could mean that our underlying notions of "logical truths" are also subject to revision. I don't think that Quine is drawing this distinction and I suspect that he means both these things. Harvey Siegel discusses the problem raised by the alterability of rational principles. He suggests

that the articulation of impartial and general principles by which we can assess reasons is part of the task of a theory of rationality. He goes on to note, however, that any such principles are linked to the traditions within which they evolve and consequently they will have different formulations at different periods in this evolutionary history. Speaking of principles, Siegel asks:

If they evolve and change how can they be impartial and universal? If impartial and universal, how can they change?
(Siegel,134)

Siegel goes on to draw an analogy between the evolutionary process of a rational tradition, the principles of assessment within that discipline, and the biological evolution of a species. He suggests that just as the selective pressures change on a species as its environment changes, bringing about changes in what are desirable features (different things are selected for), so too in the development of a rational tradition.

James makes odes which presage Quine's points on the conservatism of belief revision and the nature of our belief architecture. Pointing up why conservatism in belief revision should be the case James notes that the failure to consider the role of "older truths" has been the source of unjust criticism against pragmatism.

Their influence is absolutely controlling. Loyalty to them is the first principle - in most cases it is the only principle; for by far the most usual way of handling phenomena so novel that they would make for a serious rearrangement of our preconception is to ignore them altogether, or to abuse those who bear witness for them. (James 1980b,217)

This is not unlike one of the central points that Thomas Kuhn makes in *The Structure of Scientific Revolutions* with respect to a major revision of a conceptual schema within a scientific discipline. He suggests that when

scientists are faced with evidence or facts that tend to conflict with a scientific schema they will at first try to fit in this evidence by stretching the fabric of the schema as little as possible. When this tack fails one of the things that scientists are wont to do, Kuhn suggests, is to ignore or abuse the information that they cannot account for. It is only after evidence mounts to some great extent that the more central notions of the schema will undergo change. In essence, the schema or paradigm is replaced by some competing schema (an alternative theory or fundamental belief set) that better accommodates the new data. Up until the time that this occurs the paradigm or "older truths" continue to have an "absolutely controlling influence". Consider,

... what scientists never do when confronted by even severe and prolonged anomalies. Though they may begin to lose faith and then to consider alternatives, they do not renounce the paradigm that has led them into crisis. ... once it has achieved the status of paradigm, a scientific theory is declared invalid only if an alternate candidate is available to take its place. ... The decision to reject one paradigm is always simultaneously the decision to accept another, and the judgment leading to that decision involves the comparison of both paradigms with nature *and* with each other. (Kuhn,77)

Kuhn calls it a scientific revolution when these changes occur in elemental beliefs within a schema of a science. He opposes it to what he calls normal science, which is the general research programme that involves filling out a conceptual scheme. Normal science is the discovery of "facts" that fit into a conceptual scheme. Scientific revolutions involve a more fundamental change in view and will involve a change of what will count as fact. Furthermore, anomalous facts understood from the perspective of a new theory take on new meaning.

From within a new theory of scientific knowledge, they may instead seem very much like tautologies, statements of

situations that could not conceivably have been otherwise.
(Kuhn,78)

This understanding of the way science develops and changes is not irrelevant to our understanding of how it is that our own belief structures develop and are modified. James, Quine and Kuhn are all presenting notions of a structure of belief that involves a certain kind of internal coherence. This notion of a coherent belief structure is quite distinct from the notion of a simple pairing of beliefs with the world. As Quine notes, our belief structure only meets our experience of the world around its periphery, and our willingness to change the more central notions of the structure is impeded by a staunch conservatism. But when do we change the structure of our belief architecture, how deep do we go, and do these changes correspond with how we should be revising our beliefs? Gilbert Harman addresses these problems in his book *Change in View*.

13 Harman and Naturalized Rationality

Harman's account of rationality is distinctly naturalistic, taking note of how it is that human beings actually do reason and change their beliefs when trying to consider the question of how this ought to be done.

... normative and descriptive theories of reasoning are intimately related. For one thing, as we will see, it is hard to come up with convincing normative principles except by considering how people actually do reason, which is the province of a descriptive theory. ...any descriptive theory must involve a certain idealization, and idealization is always normative to some extent. (Harman,7)

One of the first tasks he undertakes is to distinguish between the type of reasoning that he is concerned with, which he calls "reasoned change in view" (the altering of one's beliefs) and the tenets of formal logic, which he calls "argument or proof".

Clearly, argument or proof is not at all the same sort of thing as reasoning in the sense of reasoned change in view. There is a clear difference in category. ... Consider the ... principle Modus Ponens ... Such a rule by itself says nothing at all in particular about belief revision. It may be that some principles of belief revisions refer to such principles of argument ... [but] ... rules of argument are not by themselves rule's for revising one's view. (Harman,3)

Harman highlights this point with the following example.

...Mary came to believe three inconsistent things: If she looks in the closet she will see a box of cheerios, she is looking in the closet, but she does not see a box of cheerios. Mary should not at this point infer that she does see a box of cheerios from her first two beliefs. (Harman,11-12)

The point is quite straightforward. Rules of implication in predicate logic are not the same as our evidentiary considerations in the formation and maintenance of our beliefs. On the other hand they may have some relevance and the question then becomes one of determining what relation logical principles bear to principles of belief revision.

Harman discusses the difference between two positions with which we can approach the problem of change in view, and the different types of considerations and results that these yield. These are,

... the viewpoint of our intuitions as critics and the viewpoint of our actual practice. These two approaches yield somewhat different results. As we will see, the appeal to intuition tends toward a greater degree of idealization. In particular, it tends to overlook or minimize practical limitations, such as limitations on memory or on calculative capacity. What seems wrong when these limitations are not taken into account may be quite reasonable when they are taken into account. So the two approaches can seem at least initially to yield different results. (Harman,9)

This is a point which relates to what I have called Harman's naturalism. When, in our attempts to describe what rationality should be like, we consider the capacities of human minds, we derive principles that might not even seem acceptable in the absence of these considerations.

In attempting to derive principles for belief revision Harman begins by considering what relevance logic might have to the subject. He suggests the principle that "logical inconsistency is to be avoided" (Harman,11) and a modified view, in consideration of the example above of Mary looking for the cheerios in the cupboard, of logical implication: logical closure.

Logical Closure Principle One's beliefs should be "closed under logical implication." In other words there is something wrong with one's beliefs if there is a proposition logically implied by them which one does not already believe. In that case one should either add the implied proposition or give up one of the implying beliefs. (Harman,12)

This view must be tempered, however, because it leads to all sorts of trivial and useless things being implied from one's beliefs. It is tempered by the following principle:

Clutter Avoidance One should not clutter one's mind with

trivialities. (Harman,12)

We cannot use this principle as a putative reason to refuse accepting the implications of our beliefs. It is rather intended to function as a meta-principle which,

... constrains the actual principles of revision. The principles of revision must be such that they discourage a person from cluttering up either long-term memory or short-term processing capacities with trivialities. One way to do this would be to allow one to accept a new belief P only if one has (or ought to have) an interest in whether P is true. (Harman,15)

Harman suggests that the principle of clutter avoidance presupposes that beliefs are explicitly represented in the mind. There are also implicit beliefs which are beliefs easily inferred from one's explicit beliefs.

Given that one explicitly believes the earth has exactly one sun, one can easily infer that the earth does not have two suns ... There are also cases in which one implicitly believes something that is easily inferable from one's beliefs without being strictly implied by them. An example might be one's implicit belief that elephants don't wear pyjamas in the wild. (Harman,13)

What the logical closure principle is really about, then, is the idea that we should believe implicitly those things that are logical consequences of our explicit beliefs. Even this must be tempered though. "One cannot be expected even implicitly to believe a logical consequence of one's beliefs if a complex proof would be needed to see the implication". (Harman,14)

One must first recognize an implication or an inconsistency before one has a reason to either take up or discard the relevant belief.

Harman wants to argue that belief is essentially an all or nothing proposition. We either hold a belief or we don't. He nonetheless maintains that this is compatible with belief being subject to degrees.

This is not to deny that in some ways belief is a matter of degree. For one thing implicit belief is certainly a matter of

degree, since it is a matter of how easily and automatically one can infer something from what one believes explicitly.
(Harman,22)

However, Harman also suggests that explicit belief is also subject to degrees even though one accepts these beliefs in a yes/no fashion. This is possible due to the way that rules of belief revision function.

... it may be that P is believed more strongly than Q if it would be harder to stop believing P than to stop believing Q, perhaps because it would require more of a revision of one's view to stop believing P than to stop believing Q. (Harman,22)

Explicit belief, then, is subject to degrees, but only implicitly. This is how it should be because we are not really capable of imbedding our degrees of belief in probabilistic reasoning. It is too complicated. Harman argues that probabilistic reasoning,

... leads to a combinatorial explosion, since the number of such conjunctions [needed to do probabilistic reasoning] is an exponential function of the number of possibly relevant evidence propositions. (Harman,25-26)

We need be considering probabilities for only a relatively few factors before it begins to become too much to handle.

Harman describes two competing theories of belief revision that are analogous to theories of justification, i.e., epistemological theories. These are the foundations theory and the coherence theory.

The foundations theory maintains, roughly, that some beliefs depend on others for their justification. These beliefs may depend on yet other beliefs and so on until we reach some sort of foundational level. The coherence theory does not maintain that our beliefs need to have justifications or foundations, as such, but rather that what is really essential is that they cohere. They need to be a consistent set. The coherence theory only requires justification for a belief just in case it

fails to cohere or one has some other special reason to call it into question.

Harman says that "the basic principle of the foundations theory" is that "one must keep track of one's original reason's for one's beliefs, so that one's ongoing beliefs have a justificational structure". (Harman,30) Beliefs are justified by other beliefs until we reach some foundational level. It is necessary to insist on a foundational level in order to avoid both circular and infinite justifications; infinite justifications are unacceptable for a finite being: "A belief is a basic belief if it has an intrinsic justification which does not appeal to other beliefs." (Harman,31) One should continue to believe what one does just so long as one's justifications have not been undermined. If one is left with no justifications for a belief one should give it up and also give up any beliefs for which it is the unique justification, potentially causing a chain reaction. The foundations theory contrasts with the coherence theory.

The coherence theory is conservative in a way the foundations theory is not. The coherence theory supposes one's present beliefs are justified just as they are in the absence of special reasons to change them, where changes are allowed only to the extent that they yield sufficient increases in coherence. This is a striking difference from the foundations theory. The foundations theory says one is justified in continuing to believe something only if one has a special reason to continue to accept that belief, whereas the coherence theory says that one is justified in continuing to believe something as long as one has no special reason to stop believing it. (Harman,32)

Harman says that coherence involves, "...not only consistency but also a network of relation's among one's beliefs, especially relations of implication and explanation." (Harman,32) In the coherence theory of belief revision, coherence and conservatism are the guiding principles of belief revision. These principles compete with each other.

It is as if there were two aims or tendencies of reasoned revision, to maximize coherence and to minimize change. Both tendencies are important. Without conservatism a person would be led to reduce his or her beliefs to the single Parmenidean thought that all is one. Without the tendency toward coherence we would have what Peirce (1877) called the method of tenacity, in which one holds to one's initial convictions no matter what evidence may accumulate against them. According to the coherence theory, the assessment of a challenged belief is always holistic. Whether such a belief is justified depends on how well it fits together with everything else one believes. If one's beliefs are coherent, they are mutually supporting. All one's beliefs are, in a sense, equally fundamental. In the coherence theory there are not the asymmetrical justification relations among one's ongoing beliefs that there are in the foundations theory. It can happen in the coherence theory that P is justified because of the way it coheres with Q and Q is justified because of the way it coheres with P. In the foundations theory, such a pattern is ruled out by the restriction against circular justification. But there is nothing wrong with circular justification in the coherence theory, especially if the circle is a large one! (Harman,32-33)

Harman considers the objection to the coherence theory that there are indeed asymmetrical justifications for our beliefs. He offers an example of a student, Karen, given an aptitude test. Some of the results correspond with her demonstrated abilities in the courses she has taken. Others do not. Karen, nonetheless, accepts the assessment of her aptitudes that conflict with her experience. For example, the test indicates that she has a low aptitude for history, whereas she has done well in her history course. She thus concludes that the course she took must have been an easy one.

It seems quite clear that in reaching these conclusions, Karen bases some of her beliefs on others. Her belief that the history course was easy depends for its justification on her belief that she has no aptitude for history, a belief which depends in turn for its justification on her belief that she got a low score in history on her aptitude test. There is no dependence in the other direction. For example, her belief about her aptitude test score in history is not based on her belief that she has no aptitude for history or on her belief that the history course was an easy one. (Harman,33-34)

The coherence theory maintains that while it is true that some of these beliefs lead to others in an asymmetrical or one directional way, this is merely a causal relationship. Once these beliefs have been established in this causal way they no longer depend on their causes for their continued existence.

Now the story goes that a few days later Karen is told that the results she was given were not her own and her results have been lost. According to the foundations theory she should abandon all the beliefs which were based on that information, unless she has some other justification for them. According to the coherence theory she should not. Her beliefs "acquire justification simply by being believed" and "abandoning all these beliefs is costly from the point of view of conservatism". (Harman,35)

Harman suggests that despite the intuitions of most of us that she should alter her beliefs given this controverting evidence for their justification, the psychological evidence in fact suggests that she would keep her beliefs. This psychological phenomenon is usually referred to as "belief perseverance in the face of evidential discrediting" or just simply "belief perseverance". Harman suggests that belief perseverance is not explainable by a habit theory of belief, though habit may play some role in belief.

...there are cases in which one has to struggle in order to abandon a belief one takes to be discredited. ... this is not what is happening [in belief perseverance]. Subjects ... are not struggling to abandon beliefs they see are discredited. On the contrary, the subjects do not see that the beliefs that they have acquired have been discredited. They come up with all sorts of "rationalizations" (as we say) appealing to connections with other beliefs of a sort that the coherence theory, but not the foundations theory, might approve. (Harman,38)

This is positive evidence that the way in which we actually function is more along the lines of a coherence theory than a foundations theory. Harman wants to say that it is not merely the case that we function this way and that this settles the normative question. Rather, he suggests that the foundations theory implies that we should keep track of our justifications for each of our beliefs, but this is neither what Karen did, nor is it what people normally do. Indeed if this were a requirement we would be left with very few beliefs. This is because we are limited in our capacities to remember justifications; we can't keep track of all the information that we would use to support our beliefs. We must avoid clutter. If we take this as evidence for the truth of the coherence view then, as we have already seen, it leads to a principle of conservatism (we should change as few beliefs as possible) which is counterbalanced by coherence. Here Harman suggests the following principle:

Principle of Positive Undermining One should stop believing P whenever one positively believes one's reasons for believing P are no good. (Harman,39)

He contrasts this with the corresponding principle in the foundations theory:

Principle of Negative Undermining One should stop believing P whenever one does not associate one's belief in P with an adequate justification (either intrinsic or extrinsic). (Harman,39)

He then goes on to say,

The Principle of Positive Undermining is much more plausible than the Principle of Negative Undermining. The Principle of Negative Undermining implies that, as one loses track of the justifications of one's beliefs, one should give up those beliefs. (Harman,39)

He notes that it is a matter of fact that people by and large do not keep track of the justifications of their beliefs. On the basis of this

observation, if one were to adhere to the foundations theory one would be led to the absurd conclusion that we should stop believing most of the things that we do believe. The Principle of Positive Undermining lacks this absurd conclusion.

The whole coherence schema of belief revision reaches a limit in investigations where we accept something as a working hypothesis. When we do so we function in relation to the hypothesis based on the principles of the foundations theory. (Harman,49) That is, we keep track of justifications that are unidirectional and we keep the (hypothetical) belief only so long as it promises to be fruitful. In other words, we throw out the principle of conservatism which can soon lead to clutter problems that put a limit on our ability to broaden our scope of hypotheses too far. Indeed, the promise of fruitfulness (investigative and alethic) is why we take up hypotheses.

Accepting a particular working hypothesis is fruitful if it allows one to make sense of various phenomena; if it leads to solutions of problems, particularly when there are independent checks on these solutions; and if it leads naturally to other similarly fruitful hypotheses. If a working hypothesis is sufficiently fruitful, one may become justified in fully accepting it. (Harman,46)

Harman goes on to note that we fully accept a hypothesis if it is sufficiently fruitful as to suggest that it is true. This is a difficult process and there is a tendency to "... convert tentative acceptance to full acceptance in a way that seems overly hasty to critical reflection". (Harman,47)

Full acceptance ends inquiry. There are two indicators of full acceptance: 1) One is no longer actively trying to determine the truth of something, and 2) One uses what one has accepted as a starting point for

further theoretical and practical thinking. Considering the principle of conservatism:

Full acceptance ends inquiry into P in the sense that, having accepted P, one is justified in continuing to accept P in the absence of a special reason to doubt P or at least a special reason to reopen one's inquiry. (Harman,48)

There are different sorts and levels of commitment in accepting something.

One accepts things in different ways. In addition to the distinction between fully accepting or believing something and accepting it merely as a working hypothesis, there is also a distinction between accepting something only for oneself and accepting it as a member of a group. (Harman,50)

We depend on each other for information. We each accept things on the say-so of others, and we also offer information in an authoritative way - guaranteeing the truth of it.

The use of authority in this way is important. It promotes the growth of human knowledge. Each person does not have to start over again from scratch. Appeal to authority also plays a role in learning new terminology, since one must often accept certain principles, at least as assumptions or working hypotheses, before one can use the terminology in a way that constitutes mastery of the terminology. (Harman,51)

Authoritative statements most often occur in group inquiries. Indeed they are often indicative of a component of group inquiry which is the accepting of something for the group. Harman suggests that special constraints apply to our acceptance of things for a group.

... a member of such a group should not accept something in an authoritative way for the group as a conclusion of the group unless he or she can coherently suppose there is no relevant evidence he or she hasn't considered that is possessed by others in the group. (Harman,51)

Acceptance for a community, then, is a stronger kind of acceptance. It involves guaranteeing or promising to others the truth or at least a strong evidential grounding of the information being offered.

Harman goes on to note that,

... such group inquiry is in fact the usual case. Learning about the world is a cooperative enterprise. One comes to accept things as a member of one's family or society or profession or culture. It is only when people become methodologically self-conscious that they distinguish their own private opinions from the things they accept as members of a group. (Harman,51)

Harman goes so far as to suggest that even in the event that somebody has a justified true belief for some proposition we will not say that that person has knowledge if the justification for this belief does not take into account undermining evidence that the community at large has, even though this undermining evidence is in fact false. One must, as a knower, always participate in a community and unawareness or ignorance of the full weight of relevant evidence held by the community compromises one's ability to "know". (Harman,51)¹

14 The Open Ended Discourse of Discourses

In looking at some methodological considerations for rational procedure presented in Harman the question arises whether these are enough or even the right methodological principles. As Rescher has noted, rationality has to do with intelligence and intelligent procedure. I think it would be foolish to suggest that there could be some finite set of principles that could guide us in rational or intelligent procedure. There is no rationality kit that would allow us to proceed through the world intelligently. But a key thing to note about this situation is that it does not throw the possibility of rationality to the wind. Indeed, there are infinite things that we might want to know to aid us in rational procedure.

¹ This notion of community inquiry is central to the figures I have considered above: James, Quine and Kuhn. Indeed pragmatism is the beginning of the notion of truth seeking as community based at the most fundamental level.

Rationality generally is not separable from various distinct areas of knowledge. If I understand auto mechanics I am much more capable of proceeding rationally in the repair of the car than if I do not. Analogously, understanding certain things about the workings of the human mind can be an invaluable asset when it comes to using one. Does this mean that the first thing anyone should do who wants to pursue the goal of rationality is undertake the study of psychology? Not likely. Or at least maybe not as we typically understand it. But there are some insights to be gained from experimental psychology. If we hear evidence that suggests the existence of the perseverance of belief after refuting evidence, then keeping this in mind might help us with rational procedure. There is much psychological information contained in many other places than the halls and annals of professional psychologists. There is also a great deal of information, other than psychological information, that has to do with a multitude of other things that are pertinent to our reasonings. The point really is just this: There is no finite amount of information that will allow us to proceed rationally. We can always know more, and knowledge is eminently pertinent to rational procedure. These statements, nonetheless, merit a caveat. Our abilities as human knowers are limited, and so, as Rescher would note, we must at some point make our decisions (to believe or act) based on incomplete evidence.

Furthermore, truth is only one good amongst many and it is not likely the highest good for humans. It is, rather, a good that should be subservient to other higher goods. I don't seek knowledge above all else. This is clear from my example above where I suggested that if I suspect my wife of having an affair but lack any really good evidence then I might

just leave well enough alone. But if this just sounds like willing blindness, I think it can be made clear that there are instances in which giving up the search for knowledge (truth) is not being willingly blind but is rather done in the search for a higher good. Much as the need to act and decide what to believe requires us to reach closure with incomplete, and often quite limited evidence, because we are limited beings, so our pursuit of truth (in the sense in which I have been using it in my discussion of cognitive rationality) will also be limited by our desires and needs to pursue other things that are good. When I pursue art I do not pursue knowledge in the theoretical/ cognitive sense which I have been discussing. The relationship between these other kinds of discourse in which we participate (and the sense in which they lead us to some kind of knowledge and truth) and the discourse which is our pursuit of theoretical knowledge, is another issue I can't really get into here. Wilfred Sellars, in "Empiricism and the Philosophy of Mind," suggests that philosophy (the pursuit of theoretical truth?) "... is once again directed toward that articulated and integrated vision of man-in-the-universe - or, shall I say discourse-about-man-in-all-discourse - which has traditionally been its goal." (Sellars,104) Perhaps the pursuit of theoretical truth is the discourse of all discourses, as opposed to merely one discourse amongst many. Perhaps it is only the world that can be thought of as the discourse of all discourses.

15 Meta-Cognition and Our Lack of Control

One of the principle concerns in this discussion of rationality has been meta-cognitive processes, or thinking about thinking. And more specifically, doing this in a way that yields something, which allows us to control and improve our thinking and our getting around in the world generally. Daniel Dennett in *Elbow Room* addresses the kinds of meta-level decisions that a "self-controller" makes. He suggests that for some fairly simple meta-level problems of control there are fairly straightforward answers, which he sums up in such phrases as: "look before you leap, a stitch in time saves nine, don't cry wolf." (Dennett 1984,86) We soon move beyond these simple problems, though.

Among the questions facing a sophisticated self-controlling agent are: could I revise my basic projects and goals in such a way as to improve my chances of satisfaction? Are there grand strategies or policies that are better than my current ones? Is there a style of operation that would suit my goals better than my current style? Will my current policies tend to lead me into tight quarters with little room to manoeuvre and great risk of disaster? What should my general policy be regarding risk? What kind of an agent do I want to be or become? ... There are in general no "book answers" to these questions and so, as in the mid-game in chess, one must abandon the book openings and strike boldly out into the territory of risky, heuristic reasoning. (Dennett 1984,86)

The ability to pose these meta-level questions to oneself requires the ability to suspend one's beliefs, desires and intentions in order to evaluate them objectively. "But once this examination extends beyond the mere resolution of immediately conflicting goals (for example, I can't go to the theatre and the concert at the same time) the 'weighing' can invoke an indefinitely large and unstructured set of further considerations." (Dennett 1984,86) We must limit our reflections in order to get on with it, and in doing so we are always leaving things out - even things that are

potentially important.

Since we are thus inevitably confronted with the need for a limited and incomplete survey of considerations, it becomes important to us that the "right" considerations occur to us *early* in the deliberation process. But we cannot directly control this ordering; we cannot play parade marshal for the queue of considerations-to-be-reviewed, putting each in its proper place in line. For such control takes time and effort, and we are rushed. We must relinquish control over the very process that generates the incomplete set of considerations on which we act, and hence we are always somewhat at the mercy of that process. However, knowing that we are and always will be at the mercy of that process, we can take steps to improve - if not guarantee - its reliability. (Dennett 1984,86-87)

Even though in its immediate functioning the process orders what considerations will come to mind when we are attempting to decide (resolve a problem, choose between options, etc., i.e., what will fall within the frame (of considerations)), we control that framing process in the long run. Dennett suggests that "... further levels of reflection, further bouts of self-evaluation, *tend* to lead to improvements in the 'character' of the agent" because these lead to self knowledge, and "... up to some ill-defined point, the acquisition of self-knowledge must be a good thing..." (Dennett 1984,87) However,

...even this process of reflective self-criticism and meta-level control reasoning must have its limits. If a complacent rejection ... of all opportunities for reflection would be ill-advised, it would clearly also be irrational to embark on a limitless round of self-evaluation. The unexamined life may not be worth living, but the overexamined life is nothing to write home about either. (Dennett 1984,87)

Nonetheless, we periodically venture into deep levels of evaluation (though even here they are not limitless). Evaluation is open-ended and always subject to re-evaluation. The deeper levels of evaluation are less clear because they call into question the very language that we use to describe things, and there is no meta-language within which this can be

done. This is a sort of Nietzschean revaluation of all values. Dennett regards Nietzsche's project somewhat disdainfully. Speaking of Nietzsche he says:

... it now appears that the price we pay for our open-ended capacity for reflection is entrance, on occasion, into domains where we can no longer reside any faith in our benign tendency to self-improvement. (Dennett 1984,88)

It is less obvious to me that Nietzsche has gone beyond the point of self-improvement and the rational. Quite to the contrary, while I think that the possibility of continuously getting mired in deeper and deeper levels of evaluation and never coming to any conclusions is a real danger, I don't think that it is a trap that Nietzsche has fallen into. Indeed, he revalues values in an eminently rational way. This is not your everyday sort of rationality. It is not deciding whether to go to the theatre or the concert. It is not even deciding whether to marry or not. It is even deeper than considerations of whether our justice system is really just. It gets down to the level of what we mean by the words such as good and evil. I don't think that this is irrational as a project, though if it were to paralyse an individual and make her incapable of acting - even the individual undertaking the revaluation as a philosophical project - then it would likely enter the realm of the irrational. As a philosophical project, though, it is not only quite rational, but also can prove to be very valuable.

16 Conclusion

I have tried to argue in this chapter that we must conceive of rationality as dealing with the limitations of the human mind when attempting to understand the world. There is some systemic totality to the world, but all that we can have access to is some limited portion. We function within a discourse that attempts to describe this totality but it always falls short. Indeed we never know precisely what is missing, and not knowing what is missing we know not where to look for it because relevance is usually a contingency. Rationality nonetheless enjoins us to seek truth and understanding. Correspondence is not viable as a theoretical basis on which to do this because this would require some access to reality beyond the discourse in which we participate in order to verify correspondences. This, however, is not how we check our notions of reality even with sense experience. I have argued that even our belief structures face the test of empirical information as a unified whole. All that we can do is create a unified coherent system of beliefs with which we attempt to describe reality. This is true for the kinds of epistemological reasons I have just pointed out and also true for the cognitive reasons that Harman points out.

When we understand our attempts to come to knowledge in this way we also must appreciate that it is an inherently risky endeavour. This means that the notion of "risk assessment" takes on new epistemological weight. What's more, while we cannot limit the inherent risks in our attempts at correct belief we can reduce them through the meta-cognitive processes by which we assess our own cognitive success. We can reflect on our own cognitive path. Is it working? Is there a better one? Even

this reflective analysis of our own thinking, however, must have its limit. I suggest that as a philosophical project (in the case of Nietzsche) this limit is much farther out than it is in other cases. In this chapter I have tried to show how this limiting can be achieved on a reflective conscious level in which we have recourse to a normative theory of thought or rationality. The problem of how this limiting occurs and whether it is really under our control (volition) or is more just part of our hard cognitive architecture comes up in Chapter IV. I explore there how emotion might play a role in our functional cognitive architecture and aid the process of limiting our cognitive purview at a pre-conscious level. In the next chapter I look at philosophical theories of emotion in order to help us understand what role they might play in our cognitive lives.

III PHILOSOPHICAL THEORIES OF EMOTION

1 Introduction

In this chapter I begin with a discussion of how we might assess emotions for rationality and of the historical conflict between different views of the relationship between emotion and rationality. In sections three and four I go on to examine two important historical theories of emotion in Plato and Hume. They consider how emotion relates to rationality. I then go on to examine contemporary philosophical theories of emotion. The various theories place differing emphases on various aspects of emotion (e.g., bodily disturbance, informational content, role in motivation, etc.) and not surprisingly the most acceptable view seems to be one that combines these views and tries to strike a balance in accounting for the various contributing elements of emotion. Finally, in section six, I consider two contemporary analyses of the (logical) structure of emotion as an intentional state. Through this analysis I hope to clarify the role that emotion has to play in the other aspects of our cognitive lives which I go into in Chapter V.

2 Emotion and Rationality

Can emotions be viewed as rational in their own right? De Sousa notes that there are conflicting traditions in Western thought with regard to emotion. One says that emotions are essentially irrational and a threat to rationality. He writes:

Ira brevis furor, said the Romans: anger is a brief bout of madness. A long tradition views all emotions as threats to rationality. The *crime passionel* belongs to that tradition: in some jurisdictions it is a kind of "brief-insanity defense." We still say that "passion blinds us", and in common parlance to be philosophical about life's trials is to be decently unemotional about them. Many philosophers have espoused this view,

exhorting Reason to *conquer* Passion. (de Sousa,4)

There is, however, an opposing tradition which reverses the hierarchy. Hume says, for instance, that "reason is and ought to be nothing but the slave of the passions". (Hume,415) In terms of this tradition, epitomized by Hume, emotions are not of themselves rational, but are the command center that offers directives which it is up to rationality to discover the means for.

There is nonetheless a third position in the history of western thought that argues for a stronger connection between emotion and rationality. De Sousa notes that Plato held that emotion (*thumos*) "typically sides with the rational." (de Sousa,27) It is this third position that is evident in the western tradition with which de Sousa wishes to identify and ally himself. Speaking of emotions de Sousa rejects the possibility that,

... they are like sensations, stomachaches or involuntary twitches. For if they were, they would not be amenable to rational evaluation. In fact, three sorts of considerations link emotions directly or indirectly to rationality: our confidence in judgments of *reasonableness*, the use of emotions as *excuses and justifications*, and the *thought-dependency* of most emotions. (de Sousa,5)

Now I want to look a little more closely at the contrasting positions presented by Plato and Hume.

3 Plato

Plato is the progenitor of a long history in the Western philosophical tradition of giving consideration to emotion or the passions (so called because they were considered to be passive as opposed to active) as an important aspect of mind. As I have noted above (in the section on mind and body) he divides the soul (or mind, i.e., psyche) into three parts. Thumos or the spirited element is most typically glossed as emotion, and it is contrasted with both the rational element and the appetitive element. He sees these elements as often being in conflict with each other. Consider what Socrates says in the following passage from the *Republic* in which he contrasts the spirited element with the appetitive element:

I am more inclined to put my faith in a story I once heard about Leontius, son of Aglaion. On his way up from the Piraeus outside the north wall, he noticed the bodies of some criminals lying on the ground, with the executioner standing by them. He wanted to go and look at them, but at the same time he was disgusted and tried to turn away. He struggled for some time and covered his eyes, but at last the desire was too much for him. Opening his eyes wide, he ran up to the bodies and cried, 'There you are, curse you; feast yourselves on this lovely sight!' ... The point is surely that anger is sometimes in conflict with appetite ... (*Republic* 439e-440a)

Plato imagines the soul as three independent faculties that interact, and says that it is only when the rational element is in control that there is justice. Plato's notion of independent faculties vying for power in the soul of the individual is adopted throughout much of the Western tradition. Ultimately he wants to say that just as justice in a society depends on each of the elements of society playing their proper role including the condition that the proper rulers are the ones who are actually ruling, he wants to say that justice in the individual depends on the rational part of the soul assuming the role of leader for the other aspects of soul. His

notion of rule by the rational element, nonetheless, is not that of a tyrannical rule and is more sophisticated than it is often given credit for being. Hume, as I have already noted, assumes a rather strikingly different position.

4. Hume

In *A Treatise on Human Nature* Hume writes that "reason is and ought only to be the slave of the passions". (Hume,415) This is quite a reversal on what is typically said about the relation between reason and emotion, as I noted above in the section on Plato. Why, then, does Hume want to say that reason is the slave of the passions?

In Bk II, Pt 3, Sect 3 of the *Treatise*, Hume divides reasoning into two types and goes on to see whether either of these types has any influence on the passions. "The understanding exerts itself after two different ways, as it judges from demonstration or probability; as it regards the abstract relations of our ideas, or those relations of objects, of which experience only gives us information." (Hume,413) Hume uses an analogy of a merchant who wishes to know the "sum total of his accounts" but only in order that he may know the effects of paying his debts, to argue that the first of these two types of reasonings cannot influence our actions. He writes: "Abstract or demonstrative reasoning ... never influences any of our actions, but only as it directs our judgments concerning causes and effects." (Hume,413)

Hume further on writes: " 'Tis from the prospect of pain or pleasure that the aversion or propensity arises towards any object: And these emotions extend themselves to the causes and effects of that object, as they are pointed out to us by reason and experience." (Hume,413) For

Hume then reason in no way selects or chooses which objects will attract or repel us. We are repelled, for example, from objects that will cause us pain, but we only know which objects these are through causal reasoning. Nonetheless, it is not the reasoning that causes our repulsion. It is rather through the discovery of causal reasoning that the object will cause us pain that the emotion of aversion arises as a consequence of this reasoning. He writes:

It can never concern us to know, that such objects are causes, and such others effects, if both the causes and effects be indifferent to us. Where the objects themselves do not affect us, and 'tis plain, that as reason is nothing but the discovery of this connexion, it cannot be by its means that the objects are able to affect us. (Hume,414)

Hume goes on to say that since reason does not give rise to volition, it is also quite obviously the case that reason does not oppose volition, for it would only be able to do so by "giving an impulse in a contrary direction to our passion." Having argued thus far Hume says, "Reason is and ought only to be the slave of the passions...", yet he notes that, "As this opinion may appear somewhat extraordinary, it may not be improper to confirm it by some other considerations." (Hume,415)

Hume begins these "other considerations" by saying that a passion is an "original existence" which contains no "representative quality, which renders it a copy of any other existence or modification". He is saying that passions do not refer to objects in the world and that since truth is a function of the agreement of ideas as mental copies of objects in the world, then passions can be neither true nor false. Now, reasonings are either true or false since they concern ideas and the relations of ideas, and though they may be used in the service of the passions they do not influence them. Passions are only unreasonable in as much as they may be

accompanied by a false judgment.

Passions and reason are "never in dispute for the government of the will and actions."

The moment we perceive the falsehood of any supposition, or the insufficiency of any means our passions yield to our reason without any opposition. I may desire any fruit as of an excellent relish; but whenever you convince me of my mistake, my longing ceases. (Hume,416-17)

It is simply a matter of the psychological ordering then, that for Hume reason works to discover the means to achieve the goals set out for it by the passions. He goes on to explain why it is that we become confused and think that it is reason that directs us at least at times. "Reason ... exerts itself without producing any sensible emotion." Similarly, "there are certain calm desires and tendencies, which, tho' they be real passions, produce little emotion in the mind..." So:

When any of these passions are calm, and cause no disorder in the soul, they are very readily taken for the determinations of reason, and are supposed to proceed from the same faculty with that which judges of truth and falsehood. Their nature and principles have been supposed the same, because their sensations are not evidently different. (Hume,417)

5 Theories of Emotion

When we begin to give consideration to the question of emotions one of the first things that we need to clarify is what we mean by this term. There are certainly broader and narrower ways of understanding it. Typically emotions are considered to be fairly immediate states (not to beg the Rylean, i.e., behaviourist question of whether they are in fact dispositions) that are distinguished from longer lasting states such as moods or character traits. (One way of understanding moods and character traits is as relatively long-lasting dispositions or tendencies toward certain emotional states, given certain circumstances.) William P. Alston in "Emotion and Feeling" suggests the following list of factors essential to emotion (which he derives from an example of an individual being struck with fear upon the approach of a funnel cloud):

- (1) A cognition of something as in some way desirable or undesirable.
 - (2) Feelings of certain kinds.
 - (3) Marked bodily sensations of certain kinds.
 - (4) Involuntary bodily processes and overt expressions of certain kinds.
 - (5) Tendencies to act in certain ways.
 - (6) An upset or disturbed condition of mind or body.
- (Alston, 480)

He goes on to suggest that some of the principal accounts of the meaning of emotion have to do with different emphases among these different elements of emotion. Thus he suggests that there is a strong tradition which he calls the feeling theory. This tradition,

... takes the conscious feeling to be the emotion. This view has a number of versions, differing according to the general psychological scheme in terms of which it is stated. Thus Descartes conceives of a "passion of the soul" as a "perception" (in other words a conscious state) in which the soul is passively affected, as in sense perception, but in which what is perceived is attributed to the soul itself, rather than to some physical body. For Hume, passions are "impressions of

reflection" unique kinds of experience which arise as a result of sense perceptions ... What is common to all these views is the conviction that what makes a condition an emotion, and what makes it the particular emotion it is, is the presence in consciousness of a certain felt quality, which like sensory qualities is completely accessible to introspection and accessible in no other way. (Alston,480-81)

The feeling theory, then, focuses on the felt quality of the emotion as its defining characteristic. There arise objections to defining emotions simply in virtue of their felt quality, not the least of which is the fairly widely accepted psychological truth that our own minds are not transparent to us. Alston suggests the following example: "... I might feel quite tense around a certain acquaintance; I may be aware of this simply as tenseness, without understanding its source, whereas in fact I am quite angry with him about something." (Alston,482)

De Sousa suggests that the feeling theory runs into problems when we attempt to differentiate between the kinds of feelings that emotions are and other types of feelings such as sensations or perceptions. This is especially true with respect to their objects. Many types of feelings have no objects. There is no object to my itch; I just simply itch. But, "the main deficiency of the feeling theory, however, is that there are too many aspects of emotion about which it does not tell us enough." (de Sousa,38), and for this reason we look to other theories to fill in the gaps.

Looking to fill in these gaps Alston gives consideration to other theories that also have very specific foci. He suggests that what he calls the "motivational theory" also occupies an important place in the tradition of defining emotion. The motivational theory suggests that emotions are simply part of a motivational process and occur in the process of identifying things as either desirable or undesirable and the corresponding

efforts for either acquisition or avoidance. If we refer back to Alston's list of essential factors in emotion we see that this theory focuses on two of these: "the cognition of something as desirable or undesirable", and "tendencies to act in certain ways". This theory has certain similarities to what are sometimes referred to as cognitive theories of emotion, yet is distinct in that it takes the actions that are likely to arise out of (be motivated by) emotions to be part of the emotion.

In discussing cognitive theories of emotion de Sousa says:

Cognitive theories focus on what I have called the *informational* content of emotions. This takes two forms. The more mundane emphasizes that the perception of the object of an emotion is an essential element of the emotion itself. In this sense, the objects and properties perceived are ordinary ones: to fear the lion that is coming at me, I must first detect it. But though perception in this ordinary sense can ground an emotion, it cannot *be* the emotion. The second informational element takes as its object not just the presence of the lion but its fearfulness. The emotion of fear could be the perception of that. (de Sousa,40)

According to this kind of cognitive theory of emotion, an emotion is a cognition of something as having a certain property which elicits the emotion. Cognitive theories of emotion sometimes refer to the cognitions which lead to (or are constitutive of) emotions as judgments or perceptions. I cannot merely be aware of an object without also judging (or perhaps judgment connotes something overly cerebral and I just perceive) it to be dangerous in order for it to elicit fear.

Next Alston looks to what he calls the "bodily upset theory" which he sees as focusing on the element of physiological change that coincides with emotional states. Alston remarks that proponents of the two theories previously mentioned often note that the defining characteristics they point up are also often accompanied by some bodily disturbance, that is, some

physiological change such as increased heart rate, sweating or blushing.

William James offers the most widely cited bodily disturbance or physiological theory of emotion. But James also includes an important mental component in his theory. He suggests that it is the perception of bodily disturbance which is in fact constitutive of the emotion. His theory is quite opposed to our typical conception of the relationship between our bodily "expressions" of emotions and the emotion itself as a felt quality. He suggests that we do not weep because we are sad (feel sad), but rather we are sad because we weep. He conceives of emotions as the perceptions of physiological responses which are immediate responses to the object which elicits the emotion. What he has really done is to turn the causal order of our typical conception on its head. According to the Jamesian view our perception of a certain bodily state leads to a mental state (a perception of the bodily state) rather than the perception of the object leading to a mental state which causes the bodily state. The Jamesian view, I think, arises out of the realization that many emotional responses occur extremely quickly, not unlike purely physical reflex responses. Emotions furthermore seem to inherently involve the body. Indeed James is not the first to suggest that emotions are perceptions of things going on in the body. In this respect he is preceded by both Spinoza and Descartes.

James argues for the inherently physical aspect of emotion, suggesting as evidence that we try to imagine an emotional reaction in which the bodily disturbances normally accompanying emotional reactions do not occur. He suggests that we will be left with an intellectual perception without the actual feeling. (Calhoun,10) If this sort of thought experiment

convinces us of anything it is that altered physiological states are a necessary condition for emotions. It does not, however, convince me that they are a sufficient condition or that the direction of the causal relationship between mental and physiological states goes from the latter to the former.

The other font of the Jamesian view is a general tendency toward a kind of behaviourism. Behaviourists want to focus on the observable in psychology, viz., behaviour. Alston considers such physiological disturbances as increased heart rate, sweating or blushing. He writes:

Moreover, there is a more behavioral sense in which these states constitute disturbances; when sufficiently intense, they interfere with activities which require a high degree of coordination or control. Thus Gilbert Ryle takes it as a criterion of what he calls "agitations" that it make sense to speak of them as interfering with thinking or concentrating on tasks. (Alston,482)

Alston suggests that psychologists seeking an objective meaning for the term "emotion" have focused on bodily disturbance both in the sense of outward physical manifestations and in the sense of an inability to function normally.

Behaviourists focus on two key notions in explaining the mental in terms of the physical or behaviour. These are the notions of disposition and pattern of behaviour.

On this view some emotions are just complex patterns of behaviour. No single overt act amounts to being sad. I may be crying merely to attract attention, or holding my head down just because I am shy, or listless because I am suffering from a vitamin deficiency, or dwelling on the subject of my long-lost lover only for lack of any more interesting topic. But if I do all or most of these things and others like them, then my behaviour amounts to sadness. For a behaviourist there is nothing more for sadness to be. (de Sousa,38)

An emotion under this account, then, is a complex pattern of behaviour

involving various overt behaviours. Moreover, these complexes of behaviours will only be expressed given the appropriate circumstances to trigger the disposition. These dispositions are usually described using conditional statements. For instance, to say that I dislike Joe Brown is to say that if I see Joe I will sneer, say nasty things, or try not to remain in his presence for too long. I have a disposition to do these things when presented with the object of my dislike.

De Sousa suggests that there is a more sophisticated view of the role of dispositions.

If I like someone, I may find her smoking tolerable, perhaps even sexy. But if I don't, I'll find it infuriating. Emotions on this view are *dispositional dispositions*. This captures an important feature [of emotions]: emotions control not merely behavior but also the ways that other mental events, states, and dispositions are organized. In this more subtle form, however, behaviorism has given up its crucial epistemological advantage. That advantage sprang from the rejection of unobservable explanatory mechanisms in favor of correlations between stimuli and responses. (de Sousa,39)

De Sousa says these modifications make a behaviourist perspective more richly explanatory, bringing it closer to a cognitive theory and yet leave it open to the standard criticisms against behaviourism.

The conative theory takes a certain slant on the cognitive theory.

J.P. Sartre, the principal exponent of this theory, says,

... we can conceive of what an emotion is. It is a transformation of the world ... by magic. (Sartre,58-59)

Sartre suggests that emotion is a reaction to a difficult world. Unable to transform the world in an effective way we transform our consciousness. One of the most distinguishing features of this view is that it puts emotion in the clear territory of intentional strategies as opposed to passive states. Emotions, Sartre says, are deliberate strategies for dealing with the

difficulties that we perceive (not judge) as part of the world. This perception of the world as difficult and its resulting transformation occur at the level of pre-reflective consciousness. Though Sartre says that emotion is an intentional strategy for coping with the world this is not to say that it is the kind of thing about which I deliberate and choose based on that deliberation. It is voluntary but not chosen.

Emotion, Aristotle notes in the *Nichomachean Ethics*, leads to actions which belong to the class of actions which are voluntary but not chosen. Aristotle sees actions which arise out of emotion as voluntary because they proceed from the agent (and arise out of the character of the agent) but not as chosen because we do not deliberate about them. (Aristotle, *N.E.*, 1111a 20-22, 1111b 4-6)

De Sousa suggests that what he calls the contextualist view of emotion is somewhat akin to the conative theory of emotion.

The most restrained version of the contextualist view is that only in a social context can feelings be named and talked about. Anything we talk about we have to learn to talk about. And learning requires public criteria of ascription. ...our knowledge of our emotions is not independent of the social context in which we live. ... According to more radical versions of contextualism, the very existence of emotions depends essentially, like the value of a dollar or the moves of a game, on facts about *conventional consent*. (de Sousa, 44-45)

De Sousa suggests that this view is contrary to the myth of the naturalness of emotions and their connection to our "animal nature". But also, the more extreme view suffers precisely from the problem that emotions are not purely or merely contextual (moves in a game) and "... we comfortably ascribe to ... animals: grief, revulsion, rage, boredom, fear." (de Sousa, 45) The contextual view has insights to offer into more contextually (cognitive) dependent emotions, but as a comprehensive theory

of emotion it suffers in that, "it turns its back on biology and in this way resuscitates, in a different guise, the errors of old-fashioned dualism." (de Sousa,45)

Lastly, and quite contrary to the contextualist view, one can take an evolutionary perspective in accounting for emotions. Insofar as behaviour is a product of biology, which it certainly is, evolutionary theory has things to say about how emotions came about. Emotions must, in the long run, be advantageous for the survival and propagation of the species. This, however, is an essentially different level of description and in terms of theory, I contend, will likely yield little of what I call the (logical) structure of emotions.

6 The Structure of Emotion: Causal Genesis

Moving beyond this brief consideration of some of the alternative theories of emotion it may prove valuable to undertake an analysis of the logical structure of emotions, taking our cue from the list of essential factors that Alston offers above. K.D. Irani undertakes this kind of analysis in his essay, "Understanding Emotions," where he suggests that the following is a traditional analysis of emotion:

An emotion is viewed as having the following three constituents: i) A belief state brought about by perception, memory or a combination of these. The belief is about some state of affairs. ii) This belief leads to a feeling, which is the second constituent. iii) The belief or feeling lead to a want or inclination to act in a certain way, which may or may not be actualized in action. (Irani,1,2)

Irani refers to this as the "tripartite structure" of emotion. Each of the elements is a moment in the development of a particular emotion token in which a "generative condition" gives rise to a "feeling" (about something, i.e., it is intentional), and this in turn gives rise to a desire to do

something. The relationship between these elements "is not a clear cut causal relationship." (Irani,3) Irani's focus is on emotion as a sequence of events that follow a certain pattern. It contains most of the essential elements that Alston suggests a theory of emotions should account for except that it is not clear that Irani's notion of feelings incorporates the kinds of bodily feelings that would seem to be essential for Alston. If we recall the theories that Alston and de Sousa present, Irani's theory would seem to fit most neatly into the category of a "motivational theory" insofar as emotions function as a motivation to carry out some activity. This is the element to which Irani gives emphasis. He maintains that the desires or wants that come into play in an emotional response are part of the emotion itself rather than merely its product.

Irani holds that motivation arises out of a belief or feeling, and with this suggestion seems to consolidate feeling and cognitive theories. He says that the first link in the causal chain is coming to believe something about the world. Then this belief leads to a feeling, but he is not clear about whether it is the belief or the feeling that leads to the motivation to do something. The interesting thing to note about Irani's discussion of emotion though, is that he focuses on emotion as a causal sequence.

Irani's theory, broadly speaking, suggests that an emotion is a causal sequence (loosely speaking) of events that leads from some state of the world through a cognition of this (and a feeling that this gives rise to), to a motivation or desire that may or may not be acted upon. He claims to have undertaken a phenomenological analysis to arrive at this position. He asserts about his analysis that it "... enables us to understand emotions when they arise, understand emotional behavior, and

formulate the conditions under which the behavior become incomprehensible." (Irani,19) This notion of the comprehensibility of emotional behaviour will be important when it comes to consider the sense in which we can assess emotions for rationality. The approach to emotions which gives consideration to their genesis (causal relationship to the object of the emotion) and what they in turn give rise to, is the beginning of a consideration of what I have called the structure of emotion.

7 The Structure of Emotion: Intentional Adequacy

In *The Rationality of Emotion* de Sousa presents an analysis of the ("logical") structure of emotion which he defines in terms of the relation between the emotion and its object. Insofar as emotions are cognitive (and not merely physiological) and have to do with our perception of the real world they are related to objects in the world. Sometimes the relationship is appropriate or true and at other times mistaken. There are different elements having to do with the relationship of an emotion to its object that contribute to an assessment of its appropriateness. De Sousa lists the following parameters for assessing this kind of appropriateness, which I will try to flesh out with examples afterwards:

(I) Target. The target of an emotion is a real object, typically an actual particular, to which that emotion relates. The target is that real object, if any, at which the emotion is directed.

(II) Proper target. A target of an emotion is also its proper target, if and only if it would remain unchanged by the subject's possession of full relevant knowledge.

(III) Focus, focal property and motivating aspect. Emotions having targets typically involve a focus of attention, which is the apprehension of some (real or illusory) focal property of the target. Under certain conditions, which define the standard case, the focal property is also the motivating aspect of these emotions.

(IV) First causal condition. A causal connection between the focal property and the occurrence of the emotion is a necessary condition for the former to be a motivating aspect.

(V) Corrigibility. There is no privilege of incorrigibility in the subject's access to motivating aspects.

(VI) Intelligibility condition. Motivating aspects must be rationally related to the emotion they cause, in the sense that they must constitute intelligible rationalizations for the emotion.

(VII) Second causal condition. For a focal property to be a motivating aspect, it must be an actual property of the target.

(VIII) Aim. The motivational role of emotions defines their characteristic aims and acts as a constraint on the character of each specific emotion.

(IX) Formal object. For each emotion, there is a second-order quality that must be implicitly ascribed to the motivating aspect if the emotion is to be intelligible. This essential element in the structure of each emotion is the formal object.

(X) Relation of target to formal object. The target of an emotion, where it exists, is a particular that has played a crucial role in the causation of the emotion in virtue of being apprehended as instantiating some motivating aspect. The motivating aspect in turn instantiates the formal object that defines that particular emotion.

(XI) Propositional object. If an emotion has a purely propositional object, it consists in an attitude appropriate to some fact, proposition, event, or situation type that has a property instantiating the formal object of that emotion. (de Sousa, 335-336)

Now, if I am angry with Jack then he is the (I) target of my emotion. If I am angry with him because he ran off with my girl Suzy then his running off with my girl is the (III) focus or focal property of my anger. If the fact of the matter is that it wasn't Jack who ran off with my girl, but rather Joe, then Jack is not the (II) proper target of my emotion. If I am angry with Jack because he ran off with Suzy then there is a causal connection between his running off with Suzy and my being angry with him, and so his running off with Suzy satisfies the (IV) first causal

condition and hence is the (III) motivating aspect of my emotion. I may believe that I am angry with Jack for running off with Suzy, but in fact I wanted to get rid of Suzy and I'm angry with Jack for robbing me of the feeling of power that doing so would give me; my access to the motivating aspect of my emotion is (V) corrigible. Whichever of these is the real motivating aspect of my emotion, it is (VI) intelligible as a rationalization of the emotion. In order for it to be the real motivating aspect of the emotion it must be the case that it obtains; in other words, the second causal condition (VII) states that the focal property must actually exist in order for it to be the real motivating aspect. For instance, if the real motivating aspect of the emotion is the fact that Jack robbed me of my sense of power, then it must be the case that he ran off with Suzy and robbed me of my sense of power. In order for it to make sense that I am angry at Jack because he robbed me of power, there must be something about the motivating aspect, the robbing of power, that appropriately calls for anger as a response. De Sousa suggests that this is a property of the motivating aspect which he calls the (IX) formal object. Thus, the formal object of anger is whatever anger is appropriately directed towards, perhaps wrongdoing. (Perhaps, the notion of formal object is easier to understand with an example from outside the realm of emotion. For example, the formal object of belief is truth, and the formal object of desire is the good. In any of these intentional acts of believing, wanting or emoting there is an implicit axiological judgment that the formal object of the intentional act obtains.) There is a relationship between the formal object and the target (X) which runs through the motivating aspect. Thus, in this case, the formal object of my anger, i.e., the wrongdoing, is in the

motivating aspect, i.e., the robbing me of my power, which is instantiated in Jack - the target of my anger. That is, Jack robbed me of my power and this is why I am angry with him.

Jack's running off with Suzy and robbing me of my power has made me angry. Anger typically motivates us to undertake certain types of actions which have characteristic aims (VIII). In this case I plan to take revenge on Jack by letting his new employer know about his secret criminal past, and hopefully make him lose his job.

Some emotions have non-existent targets. This may occur in cases where I have simply made a mistake. I am angry with Jack, because he ran off with Suzy, but in fact there is no Jack. Or rather, Jack is a character in a soap opera about whom Suzy was talking with Glenda on the telephone. She didn't know that I was in the next room overhearing her conversation (not that it would have mattered). In my rather paranoid state I reified the object of my anger and frustration.

There is also, however, a non-erroneous case of emotions with non-existent targets. These are cases where there is a grammatical object for my emotion but no real existing target. "No particular plays the right causal role to be the target of hoping for a promotion or for life after death. ... the only object that is mandatory for hope is propositional" (de Sousa,135) I may also have an emotion in relation to a situation type as opposed to a particular (situation) which is an emotion with a propositional object. I may fear "being attacked by a murderous fiend" or "being eaten by cannibals" without having any particular or existent object in mind which I fear.

Some emotions are more thought-dependent than others and it would

seem that the more thought-dependent emotions have more specific propositional objects. For example, grief can be dispelled quite readily when I find out that the news that my mother has died is in fact wrong. On the other hand, I may have a hope which no piece of information is necessary to create nor sufficient to dispel. In either case the emotion relates to a propositional object (or situation type) in such a way that that proposition presents with the formal object. For instance, if I am worried about Suzy running off with Jack, there is no particular or real state of affairs that I am worried about. But the (XI) propositional object, under some description, still instantiates the formal object of my worry, that is, it is worrisome.

De Sousa equates the propositional object with a situation type. He then goes on to flesh out the notion of the formal object of emotion types by associating them with certain "standardized" situation types which he calls paradigm scenarios. He suggests that we learn about emotion types, their formal objects, and the particular targets, focuses and motivating aspects which can instantiate the formal objects through example. We learn paradigms of emotion types by seeing the real world, hearing stories, reading literature, etc.

A paradigm scenario is a description of a typical case of emotions being stirred in an individual in the presence of the inciting object, target or occasion. De Sousa tells a developmental story in which he wants to consider how a repertoire of emotions is built up.

We are made familiar with the vocabulary of emotion by association with paradigm scenarios. These are drawn first from our daily life as small children and later reinforced by the stories, art, and culture to which we are exposed. Later still, in literate cultures, they are supplemented and refined by literature. Paradigm scenarios involve two aspects: first, a

situation type providing the characteristic objects of the specific emotion type ..., and second, a set of characteristic or "normal" responses to the situation, where normality is first a biological matter and then very quickly becomes a cultural one. (de Sousa,182)

Children are taught to identify their reactions to certain things with certain tags (words) that identify emotions. They are in this way taught to feel certain emotions. We ultimately adjudge the rationality of emotions in terms of how closely they relate to the paradigm scenarios of the given emotion, because it is the paradigm scenario which defines the formal object for any given emotion.

Thus we assess the rationality of an emotion in terms of its success in attaining its formal object. De Sousa suggests that we call the success of an emotion in attaining its formal object "appropriateness". (de Sousa,185)

... the view of formal objects I have been expounding suggests a strict criterion for the limited category of representational states: A group of such states form a natural class, if and only if they have the same formal object. Appropriateness is the genus of the formal object in question. But that genus includes truth and goodness too. It tells us nothing about the differentiae of success conditions for emotions. (de Sousa,185)

Thus, the degree to which any representational state attains its formal object is its degree of appropriateness. For example, if belief attains its formal object then it is appropriate and we call it true. Emotions are not members of a class with the same formal object and thus must have their appropriateness determined not as a class but in each case of an emotion type.

De Sousa offers a definition of appropriateness for emotion: "... an emotion is appropriate, in part, if the target actually has a focal property in virtue of which the formal object fits the target." (de Sousa,122) That

is, for example, if I fear the approaching dark figure because I know that he wants to murder me (the object has the focal property of being dangerous which incites my fear), then my fear is appropriate because the target (dark figure) has a focal property (dangerousness) in virtue of which the formal object (fearfulness) fits the target.

What objectively in a situation warrants a given emotion? Why might it not be the case that emotional responses are always just purely the subjective responses of the emoter (in some stronger sense of "subjective" than merely belonging to that person's phenomenology)? By the way, emotions may be subconscious at times and thus (possibly) escape phenomenology. De Sousa suggests that,

The ways in which emotions are subjective do not sufficiently undermine the analogy of perception to exclude a significant claim to objectivity. (de Sousa,149)

He later goes on to consider the relation between perceiving subject and perceived object.

The question we need to ask is this: To what degree is the perceptual state ... covariant with the subject, and to what extent is it covariant with the object. ... Variations in your sensations of sight may result from (1) variations in (some sub-set of) the intrinsic or primary properties of the object, but also from (2) environmental factors (conditions of light, temperature, and so forth), (3) your physiological apparatus and its condition, (4) your experiential history and current beliefs and desires, and even (5) social and ideological factors.

Insofar as your sensory apparatus works properly, it will be decreasingly sensitive to factors (2)-(5) in roughly that order. ... this makes for an essential distinction between perception and general cognition: by and large our beliefs are modifiable in the light of factors of type (4) and (5), but our perceptions are not. (de Sousa,152)

De Sousa suggests that this distinction can be at least partially accounted for by the notion of informational encapsulation that Fodor puts forth in *The Modularity of Mind*, and later restates as follows:

By a modular cognitive faculty I mean - for present purposes - one that is 'informationally encapsulated'. By an informationally encapsulated cognitive faculty, I mean one that has access, in the course of its computations, to less than all of the information at the disposal of the organism whose cognitive faculty it is. So, for example, I think that the persistence of the Muller-Lyer illusion in spite of one's knowledge that it is an illusion strongly suggests that some of the cognitive mechanisms that mediate visual size perception must be informationally encapsulated. You know perfectly well that the lines are the same length, yet it continues to appear to you that they are not. It would seem to follow that some of what you know perfectly well is inaccessible to the cognitive mechanisms that are determining the appearances. If this is the right diagnosis, then it follows that some of those mechanisms are informationally encapsulated. (Fodor 1987,139)

Now de Sousa wants to suggest that it is in virtue of this kind of informational encapsulation that sensory perceptions as informationally encapsulated cognitive structures enjoy a decreased level of sensitivity to certain other cognitive factors which are listed above. The situation for emotion is not so fortunate.

The consequence of having no organs of emotion can now be seen more clearly: emotions are typically susceptible to the whole gamut of factors ranging from (1) to (5). At first sight this would certainly seem to make the case for emotional subjectivity ... Not only will emotions be more subjective than perceptions. They will also apparently be more subjective than other unencapsulated states - namely, beliefs. The reason goes back to some facts about physiology ... it is arguably a defining characteristic of emotion that it involves a more conspicuous participation of the body than do other mental states. Moreover, the bodily states involved tend to be part of the "slow" rather than the "fast" systems of internal information processing: those involving the relatively sedate messaging of hormones rather than nervous fibers. The latter fact acts as a kind of drag on the covariance with the properties of type (1) that we demand of a perceptual system. (de Sousa,153)

Thus, de Sousa wants to say that objectivity of emotion is analogous to the objectivity of perception. However, it is susceptible to certain kinds of factors which diminish its objectivity due to the fact that it is not informationally encapsulated as perception is. The objectivity of emotion is

further diminished by the inertia it suffers in virtue of its stronger link to the slow kinds of information processing that are part of bodily states. My body's attention outlasts the moment of my mind.

8 Conclusion

De Sousa's arguments that emotions are intentional (and essentially cognitive) states and that as such they are an integral part of our cognitive architecture gets expanded in the next chapter where I consider what role they play in possibly limiting our evidential consideration and helping to solve certain cognitive binds.

IV The Role of Emotions in Rational Cognition

1 Introduction

What will occur to me? What will I think about? Do I choose these things? Are they a product of my will (under my control), or outside my control? In addressing the question of what supplements rationality in determining our cognitive abilities, part of what we are doing is asking what supplements (or supersedes) the will and makes possible the kind of cognition of which we are capable.

De Sousa argues that it is in fact emotion which helps to overcome a certain kind of cognitive problem, viz. the frame problem, with which rational cognition is faced. Emotion does this by determining patterns of salience, and in this way, mimicking informational encapsulation, despite the fact that emotion suffers from not being informationally encapsulated. De Sousa moves from the ability to mimic encapsulation and determine patterns of salience, and its potential to solve the Buridan's ass problem, to the frame problem. I think this move is potentially problematic. De Sousa's argument involves a conjecture about what rationality would be like in the absence of emotion. This conjecture leads into the relationship between rationality and freedom of the will, which I discuss below before moving more directly into the frame problem.

2 The Determinism of Reasons and Buridan's Ass

Emotion plays a role in rationality in determining patterns of salience that draw our attention to certain things and create focuses for our attention. In this way, de Sousa says, emotion keeps us from getting stuck in certain kinds of cognitive binds, that we would get into in the absence of emotion. Thus emotion is an essential and integral part of rationality. He argues for this based on a view of what the rational mind would be like in the absence of emotion. De Sousa imagines it as rationality with something lacking, calling it, nonetheless, perfect rationality in the tradition that sees emotion as a polluter of rationality.

I think that there are problems with his description of rationality in the absence of emotion that weaken his argument for the centrality of the role emotion is to play in rationality. De Sousa imagines that rationality has to do with some kind of deductive or algorithmic process. This is not what rationality would look like in the absence of emotion if we can imagine such a thing. Indeed, I think, de Sousa's view of rationality in the absence of emotion is perhaps indicative of one of the key things that leads to the mistaken notion that there is a chasm between emotion and rationality. Rationality sounds like a clumsy algorithmic process taking one step at a time; following procedure until procedure runs out and then getting utterly stuck. Along comes winged emotion to the rescue; the very essence of fluidity and motion moving rationality to the next flow chart when it is stuck at the bottom of one and can't see the next one. But let's look at his argument for the essential role emotion is to play in rationality.

Speaking of angels, de Sousa writes:

... angels must be some sort of Kantian Rational Will, or perhaps like Spinoza's God, with goals that follow from the principles of reason itself and with a perfectly determinate system for fixing on means. By an angel, I shall henceforth mean a Kantian monster of this sort. (de Sousa,14)

De Sousa goes on to consider the problem that the angel might face if something were "left undetermined by the principles of reason". The angel, de Sousa suggests, unlike the ass stuck between two equidistant carrots, cannot be swayed by some "nonrational contingency". The angel is stuck in the following dilemma, according to de Sousa, for which the notion of free will is merely a ruse to attempt to evade the problem.

Either (1) the free decision is determined by something, or (2) it is determined by absolutely nothing. In case (2) it is simply a form of irrationality. But in case (1) then either (a) it is determined by nonrational principles, which contravenes the notion that we are dealing with a perfectly rational being, or (b) it is determined by rational principles, which contravenes the assumption that free will escapes the determination of reason. (de Sousa,15)

De Sousa goes on to suggest that we can conceive of an emotionless being as similar to the angel, though we may wish to conceive of some emotionless lower animal (e.g., an ant) as being machine-like and impelled by tropisms rather than being guided by a pure rational will. What these two beings have in common is that they are "both completely deterministic" according to de Sousa. I think that de Sousa, here, betrays a notion of rationality that is overly determined which leads to his notion of rationality as being deterministic or excluding free action.

On the contrary, rationality is neither algorithmic nor deterministic. Rationality can only be assessed of agents in the world. Thus it is never the case that rationality is "all things considered". Emotion is not the force that liberates us from our own rationality, as de Sousa seems to want to say. Nor is rationality that which keeps us from enslavement to the

passions, as a certain reading of Aristotle would seem to suggest. Insofar as we can distinguish between these different aspects of mind they are complementary - not combative. Cognitive dissonance does not require the positing of a battle between these different elements, Plato notwithstanding.

De Sousa suggests that one of the things that emotion can do in creating salience is resolve the kind of problem he suggests above where his fictitious pure rational being becomes stuck when he finds himself equidistant between equally desirable things that can be achieved with equal ease. To resolve this stuckness he considers three alternatives for making the choice. (1a) "It is determined by non-rational principles, which contravenes the notion that we are dealing with a perfectly rational being." (1b) "It is determined by rational principles, which contravenes the assumption that free will escapes the determination of reason." (2) "It is determined by absolutely nothing" and hence "it is simply a form of irrationality."

Option (2) would seem to mean that it is an event without a cause (understanding with Davidson that causes of human actions are reasons). The creature is caused to become unstuck by literally nothing. Surely something (and not nothing) moves the creature out of being stuck, whether it is forgetfulness or some other non-rational contingency. Perhaps "non-rational contingency" is intended to be covered by (1a) though it seems to me that these are distinct, since a non-rational contingency is not really a non-rational "principle," and that de Sousa has omitted one of the logical possibilities. (1a) suggests that the creature might be moved by some "non-rational principle" as opposed to a mere contingency. Thus a creature of such "perfect" rationality might have a

principle something like: "When caught in the ass's bind the item to the left will be assigned heads, the item to the right tails, and the decision will be based on a single coin toss." This "non-rational" principle could be modified by a perfectly rational being to cover possible situations which I may have not foreseen.

Option (1b) suggests that the decision could be based on a rational principle, but then the creature would not be free. I think that de Sousa is confused here both about the nature of rationality and its relationship to freedom. For instance, the non-rational principle I have suggested above is not really non-rational. If rationality is about intelligent procedure and getting around in the world (which it is), then a rational creature that found itself susceptible to becoming stuck in this way (and recognized that it was stuck when it was stuck) would certainly try to devise such a principle to get it out of these situations. In fact any principle that would suffice to get the stuck agent out of the bind would be a rational principle. Since any principle that would do the job would be a rational principle then no non-rational principle would do the job.

The objection could be raised, that the principle "Kill all blue eyed people shorter than 5' 3", would not be rational. Following the suggestions of Rescher we need to take evaluations (including moral valuations) into the fold of a more holistic rationality. This principle would thus be irrational because it is morally unacceptable. It would also run into problems from a strategic point of view. There is, I think, a problem with this counterexample, which is that this principle is supposed to be one that would get us out of the above described bind, and I don't really see how it would do that. I, nonetheless, must modify my position to say that any

principle that does not fail on evaluative or strategic grounds and still manages to get us out of the bind will be a rational principle. The point is that a rational principle will be any principle that meets the pragmatic criterion of working and no other more stringent or abstract criteria.

De Sousa expresses the "assumption" that free will must escape "the determination of reasons". There is, in fact, no such thing as the determination of reasons. De Sousa says, quite rightly, that god "needs no principles of rationality" (de Sousa, 198). He suggests however that an angel would need principles of rationality. It lacks the full information that god has but might have a perfect set of rationality principles. This makes no sense. There is no finite set of principles that make up "rationality" through which we run information to produce rational results. The having of information is an integral part of rationality, and as I argued in Chapter III, sect. 12, principles of rationality or logic are in important ways just like other kinds of information. These truths are ultimately contingent. Rationality is about getting around in the world and for this we need information about the world; it is a key ingredient of rationality. Thus, since god has no need for rationality, there is no perfectly rational being.

Oddly enough one of the objections to this situation - no perfectly rational (determined) being - is that it implies the sacrifice of freedom. The argument is roughly that every event is either causally determined or random. If rationality isn't determinate, then our choices must be random. But if our choices are merely random then how can they be ours, and how can they be free? Dennett suggests that in fact things are not either causally determined or random, but rather they are either causally

determined or not causally determined.

... if we could somehow *make sense* of human actions at the level of intentional explanation, then in spite of the fact that those actions might be physically undetermined, they would not be random. (Dennett 1978b,288)

These actions would still be intelligible. Thus despite the fact that they are not causally determined they do not suffer from the kind of randomness that would rob us of our freedom. Dennett thinks that indeed, our actions typically are undetermined. He says that there are many different ways in which any given human action can be realized and that we usually don't care which particular physical motion brings about the act we intend.

... let us suppose that our nervous system is so constructed and designed that whenever in the implementation of an intention, our control system is faced with two or more options with regard to which we are non-partisan, a purely undetermined tie-breaking "choice" is made. There you are at the supermarket, wanting a can of Campbell's Tomato Soup, and faced with an array of several hundred cans of Campbell's Tomato Soup, all roughly equidistant from your hands. What to do? Before you even waste time and energy pondering this trivial problem, let us suppose, a perfectly random factor determines which can your hand reaches out for. This is of course simply a variation on the ancient theme of Buridan's ass ... (Dennett 1978b,291)

The point is simply that this kind of action is undetermined, or it at least may be, and yet this kind of randomness in the causes of my actions does not render them unintelligible. But it was the kind of randomness that rendered our actions unintelligible that threatened our freedom.

... we can indeed install indeterminism in the internal causal chains affecting human behaviour at the macroscopic level while preserving the intelligibility of practical deliberation ... we need not fear that macroscopic indeterminism in human behaviour would of necessity rob our lives of intelligibility by producing chaos. Thus philosophers such as Ayer and Hobart, who argue that free will requires determinism, must be wrong. (Dennett 1978b,292)

Dennett suggests that there may be a more important sense in which indeterminism gets into the decision process. This kind of indeterminism, the Campbell Soup variety, he finds quite unsatisfactory.

...it seems all we have done is install indeterminism in a harmless place by installing it in an irrelevant place. The libertarian would not be relieved to learn that although his decision to murder his neighbour was quite determined, the style and trajectory of the death blow was not. Clearly, what the libertarian has in mind is indeterminism at some earlier point, prior to the ultimate decision or formation of intention ... (Dennett 1978b,292-293)

The more substantive kind of indeterminism that Dennett has in mind has to do with the kind of exploitation that intelligence makes of randomness in the process of invention. He takes his idea from the French poet Valery who suggests that invention is a twofold process wherein the first part of the process is a sort of random combination of things. The latter part is the recognition and selection of things from this mass of possibility. It is this latter part that Valery thinks is what makes up genius. Dennett attempts to apply this notion of invention to human decision making.

An interesting feature of most important human decision making is that it is made under time pressure. Even if there are, on occasion, algorithmic decision procedures giving guaranteed optimal solutions to our problems, and even if these decision procedures are in principle available to us, we may not have time or energy to utilize them. (Dennett 1978b,293)

He suggests that the process of "generate and select", which Valery applied to invention, also likely applies to decision making. I survey various different considerations. "Some of these considerations, we may suppose, are determined to be generated, but others may be non-deterministically generated." (Dennett 1978b,293) Even with an understanding of the agent's beliefs and desires we could not predict the

decision unless we knew which considerations occurred to the agent, except perhaps conditionally. We might say that if such and such considerations occur to the agent, she will make this decision, but that if this other consideration occurs she will make this other decision.

Notice that although we are supposing that the decision in this way is strictly unpredictable except conditionally by the intentionalist, whichever choice Jones makes is retrospectively intelligible. (Dennett 1978b,294)

The choice is intelligible based on which particular considerations occur to Jones. But it is precisely with which considerations that will occur to Jones where the indeterminacy of the process takes place. Whether or not this generation process is in fact physically undetermined or just patternless in all likelihood makes little difference.

What is the upshot of all this? Dennett is suggesting a way in which we can conceive of a rationality which is indeterminate (and yet fully rational), and which allows the possibility of human freedom. What's more he does this without reference to emotion. De Sousa also introduces indeterminacy into rationality by suggesting that emotion plays the role of introducing patterns of salience. Now there are a couple different ways of viewing the relation between these two views. One way is to view them as competing schemes for introducing indeterminacy in rationality. This may well be the case but they are trying to achieve different ends. Whereas Dennett wants to open up the possibility of freedom, De Sousa wants to make sure that rationality (cognition) will function unencumbered by determinacy. Despite these different ends I think that Dennett's way of introducing indeterminacy shows that emotion may not be necessary to overcome determinacy for rational functioning.

De Sousa's argument that emotion is necessary for rationality falters

because the categories that de Sousa has outlined for the understanding of the rational being are dependent on a contentious view of what rationality would be like in its absence which I hope to have shown in considering Dennett's view of the indeterminacy of rationality.

I have attempted to argue that rationality is not deterministic even in the absence of emotion. I doubt that emotion is the answer to tropisms; likely greater cognitive ability is, and it is quite likely that emotions can fit neatly into tropisms. I imagine that it is possible to have a rather determined response pattern for a certain type of situation that includes all the relevant typifiers of an emotional response.

Nonetheless, de Sousa is not totally off the mark either. I think that he is right that emotion has the ability to create salience. However, to suggest that it is this ability to create salience which overcomes the "determination of reasons" is overstating the point.

3 Emotion and Saliency

Now, de Sousa also thinks that emotion in its ability to create saliency helps to overcome certain kinds of cognitive problems, viz., the frame problem. The frame problem concerns how an information processing system (including minds), after some particular change in its data base, updates the relevant information in the rest of the data base without running through the extremely inefficient process of checking the whole data base. The system, then, needs to be able to focus in on what effects some given change will have on the rest of the system without checking the rest of the system. It needs to have access to relevancy in an a priori way it would seem. De Sousa's speculation is that in the human mind emotion helps to achieve this by creating saliency and drawing attention, as it were, to particular objects.

Distinct from the frame problem is the Buridan's ass or Campbell Soup problem, which concerns the creation of saliency in objects external to us. I'm not sure if there is any important relationship between this problem and the frame problem. What's more, I think that while de Sousa sees emotion as solving the Buridan's ass problem (and makes a leap from this to the frame problem) it is problematic what role emotion has to play in solving this problem. I think that emotion may indeed create saliency among objects, but it is certainly not the only thing or likely even the key thing that does so. But let's consider a little more closely de Sousa's argument for the role he sees emotion playing in rationality.

De Sousa suggests that emotion must play the role of determining saliency in rationality because "... there is no such thing as fully determinate rationality." (de Sousa, 1991) The central problem here is that,

"No logic determines salience; what to notice, what to attend to, what to inquire about." (de Sousa,191) As we gather information in the world, forming our beliefs, some things stand out more than others, and there are different weightings that we can give things for purposes of induction. De Sousa suggests that it is emotion which fills in these gaps, not simply by playing the role of determining our desires, but by creating patterns of salience. He states his hypothesis thus:

Emotions are species of determinate patterns of salience among objects of attention, lines of inquiry, and inferential strategies. (de Sousa,196)

De Sousa then goes on to consider examples that illustrate the functioning of emotion as a determining of salience that we encounter in our everyday lives. One such example has to do with the recognition of the immanence of our own deaths. Our emotions seem to influence our reasoning in such a way that they either guide or distort our reasoning, depending on the view we take of the appropriateness of this influence.

Consider for example the way that we are wont to discount experience for distance in space, time, or affection: we care less about pain if it is far enough in the past or future, and we are relatively indifferent to distant disasters. This seems reasonable, within vague bounds: but there are no rules of rationality that prescribe the rate of discount. The thought of our own death is subject to this discount. We are usually curiously indifferent to it, but as Tolstoy shows in *The Death of Ivan Illyich*, the prospect of death might provoke such a shift in a man's pattern of salience that an entire life is seen aright for the first time: Ivan Illyich's emotions are axiologically adequate only when he faces death. (de Sousa,197)

De Sousa goes on to consider how it is that emotions influence salience and our focus of attention. He suggests that often we have no reasons to offer for our shifts in attention; these are the results of emotions determining our shifts of attention, not vice-versa. Furthermore,

we may not be able to offer reasons for our shifts of attention. We sometimes account for this inability to offer reasons for our shifts in emotion, and consequent shifts in attention, by saying that they are somatic phenomena. De Sousa believes that his model offers a better basis for explaining this:

On my view, emotions set the agenda for beliefs and desires: we might say they ask the questions that judgment answers with beliefs and evaluate the prospects to which desire may or may not respond. ... questions have much to do with the determination of answers... In this way emotions can be said to be judgments, in the sense that they are what we see the world "in terms of." ... Much the same reasons motivate their assimilation to desire, of which one kind or component, as I have argued, can itself be classified as an emotion. Logic leaves gaps. So as long as we presuppose some basic or preexisting desires, the directive power of "motivation" belongs to what controls attention, salience, and inference strategies preferred.

For this reason emotions are often described as guiding the process of reasoning - or distorting them, depending on the describer's assessment of their appropriateness. (de Sousa,196-7)

De Sousa wants to consider the different ways in which different emotions function as determiners of our attention. Some emotions are longer lasting. We distinguish amongst, emotions, moods, and character traits in their duration as determiners of salience. This relationship is not unidirectional in nature. It is quite possible for my reasoning and my concerted attention to certain things to change my emotions. My emotions are not the only things that determine my patterns of attention.

Emotions are, in part, patterns of attention. Therefore one might expect a change in patterns of attention to entail a change in emotion. I can't be very angry any more, if I notice none of your misdeeds, nor even infer to any when there is a doubt to be resolved. (de Sousa,243)

One of the primary problems that de Sousa's example illustrates well is that it is easier to draw attention to something than it is to take it away. We

often require a distraction in order to remove our attention from something. We cannot simply stop paying attention to the object of attention unless we have simply lost interest in it. This latter case would not be a case of rationally redirecting our attention to change emotion but simply the case already described of emotion determining our attention.

It does me no good to tell myself how foolish I am to miss her: for the thought is an enemy agent, as it were, calculated to fix my thoughts on just what I should forget. I should forget her smile, her eyes, her perfect nose job The best course is to fall in love with someone else: "It'll take my mind off her." Or failing that, to hate her: directing my attention onto her betrayal, her levity, her heartlessness And the same goes for self-love: "Il est facile de se hair," said Bernanos, "Le difficile est de s'oublier." (de Sousa,243)

It would seem then that our attention is not subject to our will, and neither are our emotions which guide our attention. In essence, our ability to reason is not itself under our direct control. Rather there are certain cognitive processes which in themselves are rational cognition with no higher order homunculus to guide reason.

4 Rationality and the Frame Problem

De Sousa argues that emotion solves certain kinds of cognitive problems by mimicking modular cognitive faculties and informational encapsulation. An informationally encapsulated cognitive faculty is "one that has access, in the course of its computations, to less than all of the information at the disposal of the organism whose cognitive faculty it is." (Fodor 1987,139) I want to discuss in more detail Fodor's arguments for the role informational encapsulation has to play in making rational cognition possible.

Modular or informationally encapsulated systems do not encounter the frame problem because they have limited data bases within which they can

function with algorithmic principles for finding exactly what is appropriate to any "reasoning" within the module. A modular faculty may occur within an open or unencapsulated system and this modular faculty will not face the frame problem. This is a sort of limited solution to the frame problem by lapsing into an arbitrary delimitation of evidence and irrationality. Perhaps an open system could be composed of very many overlapping modular faculties. The fact that they overlap would make them a whole. The rule that says that any given faculty will not "gather evidence" from more than its contiguous faculties would solve the problem of reaching conclusions and fixing beliefs. This speculation on functional architecture, however, is not Fodor's solution to the frame problem. To the contrary, he thinks that encapsulated faculties are irrational and that really smart systems are unencapsulated. (Fodor 1987,141) He says:

There are, it seems to me, two interesting ideas about modularity. The first is the idea that some of our cognitive faculties are modular. The second is the idea that some of our cognitive faculties are not. (Fodor 1987,139)

His example of a cognitive faculty which seems to be clearly modular, which I cited above, is vision, or "some of the cognitive mechanisms which mediate visual size perception". He gives an example of how a certain visual illusion persists despite the fact that we have the information that it is in fact an illusion. He notes:

It's worth emphasizing a sense in which modular cognitive processing is ipso facto irrational. After all, by definition modular processing means arriving at conclusions by attending to less than all of the evidence that is relevant and available. And ignoring relevant and available evidence is, notoriously, a technique of belief fixation that will get you into trouble in the long run. ... But ... rational processes have their debilities too ... If, for example, you undertake to consider a nonarbitrary sample of the available and relevant evidence before you opt for a belief, you have the problem of when the evidence you have looked at is enough. You have, that is to say, Hamlet's

problem: when to stop thinking.

The frame problem is just Hamlet's problem viewed from an engineer's perspective. (Fodor 1987,139-140)

The challenge then, according to Fodor, is to come up with a device that is rational (unencapsulated) and yet manages to "... succeed in fixing a belief or two from time to time." You do not want to have arbitrarily delimited evidence searches and yet you want the evidence searches to come to an end. The limitations that encapsulation places are arbitrary and hence irrational but are also indicators of rationality.

Outbreaks of the frame problem are symptoms of rational processing; if you're looking at a system that has the frame problem, you can assume that it's rational at least to the extent of being unencapsulated. (Fodor 1987,141)

There are certain kinds of cognitive problems that are fairly circumscribed and which do not require, indeed avoid, access to a broader body of knowledge. "In parsing, for example, we find computational mechanism with access only to the acoustics of the input and the body of 'background information' that can be formulated in a certain kind of grammar." (Fodor 1987,141) Fodor thinks that for this reason there is no frame problem within circumscribed kinds of cognitive problems like parsing or playing chess because the relevance of information is finite. On the other hand to solve problems that require true intelligence we must have a system that is non-modular and we come upon the frame problem immediately.

Encapsulation, then, Fodor wants to say, is inherently irrational but is nonetheless an indicator of a rational system and aids the system in actually reaching the point of fixing beliefs. Fodor is certainly right that insofar as encapsulation represents an arbitrary delimitation of evidential admission it is irrational, but it is no more irrational than allowing infinite admissibility. It would seem that there are certain hard wired heuristic

devices for reaching a balance of risks for belief fixation. Our very cognitive makeup limits investigation and moves us on to belief fixation, balancing the risks of believing wrongly and not believing at all - perhaps equally dangerous situations. Encapsulation is not, however, for Fodor, the solution to the frame problem either in Artificial Intelligence (AI) or in understanding the functioning of the human mind. Indeed, he has no solution, and presents arguments against what he sees as the several pseudo-solutions that are offered by AI researchers. He thinks that they are missing the crux of the issue.

I reiterate the main point: The frame problem and the problem of formalizing our intuitions about inductive relevance are, in every important respect, THE SAME THING. (Fodor 1987,148)

We are incapable of formalizing relevance in induction precisely because we cannot determine relevance in induction beforehand. If I am engaged in inductive reasoning I cannot pick out in an a priori way all the categories of evidence that might be relevant. There is, it very often seems, nothing that methodical about how I get to the evidence that I eventually decide is in fact relevant. It may be the case that relevance is always just a question of degree and that the process of trying to obtain what is relevant must ultimately revert to the question of risk assessment which I considered above.

5 Emotion and Encapsulation

Is emotion one of these meta-cognitive devices that delimits the consideration of evidence and moves us on to belief fixation? Emotion is almost certainly not an encapsulated faculty in the way in which visual perception is. Nor is it, it would seem, the sort of cognitive function that is encapsulated by dint of having inherently circumscribed parameters of

relevance in the way parsing and chess playing do. How then, does de Sousa's speculation that emotion mimics encapsulation work itself out, if at all?

I think that if emotion is to play the role of creating salience we must conceive of it less as a part of the functional architecture of the mind or perhaps not so much as a part of the "hard" architecture. I say this because in the case of the encapsulation of the faculty of vision the evidence for encapsulation is that the functioning of visual perception is unaffected by relevant information when that information is of the wrong sort (propositional) and hence becomes "inadmissible" on architectural grounds. In parsing and chess the limits are conceptual; the types of things which might be relevant are simply very limited. Now neither of these situations obtains with respect to emotion. It is not in itself an encapsulated faculty.

There is a wide variety of types of information which can be relevant to emotion and what's more anything that is relevant seems to be "admissible". There are admissibility problems which arise in that during an emotional response other kinds of information or evidence can seem to be shut out. I would argue that this is not encapsulation, however, because the evidence is shut out only on a temporary basis and when a strong emotional response begins to subside the channels open up again. What's more this kind of mimicking of encapsulation which an emotional response achieves is a product of the character of the individual where we understand by character here simply the tendency to have certain kinds of emotional responses, etc. While the emotional response mimics encapsulation as it is occurring and shuts out other information, it nonetheless remains a

product of character and this can be influenced and changed on a long term basis (at least to some extent). Therefore, this is not encapsulation in the same way that the other two types of encapsulation are.

Nonetheless, it would seem to be able to mimic encapsulation on a short term basis and in this way influence salience by limiting "admissibility".

The key benefit of encapsulation is that in its limiting the parameters of evidence it allows us to terminate deliberation, or some more fundamental cognitive function that is a correlate of deliberation. Emotion, by mimicking encapsulation limits our consideration of relevant evidence. "Love makes you blind." Other emotions might certainly focus our attention on specific types of things (characteristics of situations or individuals), as I have discussed above. This is, of course, a double edged sword. It can help bring deliberation to a close, but closure is sometimes premature. According to Fodor this arbitrary limitation of considerations is inherently irrational. I think that this is too strong, at least for the sociobiologist's reason (which I outline below). According to de Sousa this kind of closure is essential to rationality. As I have expressed above I doubt that this is the case. That is, I think greater cognitive ability implementing principled pragmatic limitations is the key solution to Hamlet's problem and not a more rigid way of bringing about closure.

Emotion does have an essential role in rationality, besides its contribution to salience. Rationality is (in every interesting case to date) a product of human minds. Emotion is also an integral part of human beings and as such it must be brought into the fold of reason as must other things that some unreasonable people would rather leave aside, e.g., ethics or the fact that we are bodies.

The sociobiologist could also put forth the argument that emotion's ability to speed up the deliberation process may be the very key to survival in certain situations and the existence of these emotional responses is the evidence in support of this argument. For instance, my fear makes me flee the perceived danger before I look for further evidence as to whether my perception of fear is justified. This kind of accelerated process can be very useful for its survival value. From the sociobiologist's point of view, the proof that it works is in the fact that we have these emotional responses. The process of biological evolution has led us to have them because they work.

V Conclusion

The story I've tried to tell here deals with certain problems presented by an infinite universe. If the universe is what that name implies, viz., a complete system, then that would seem to offer one *prima facie* reason for us to suspect that determinism might be true. But if the universe is infinite how would we ever know that determinism is true? Indeed, now would we know that anything at all is true? In some sense everything is relevant to determining the truth of everything else. What we are charged with doing as finite "knowers" is setting some arbitrary limits on relevance that are less than infinite. Indeed what I have tried to argue here is that this notion of setting a limit is central to rationality. What this means epistemologically is that the sceptic in some sense is right because our conclusions are almost always flawed in some way. This is because there is a flaw in the methodology whereby we arbitrarily limit investigation and reach conclusions. The sceptic is wrong, however, if she wants to suggest that this means that we should give up investigation. Ultimately our investigations participate in a discourse that serves us independently of its truth value. This, however, is not to say that our limits should be made willy-nilly. Using pragmatic guidelines we must strive to set those arbitrary limits in the least arbitrary way possible. We need to be principled in our determination of these limits.

I argued that in setting out the normative guide of how this is to work we must consider what is in our cognitive nature. The latter section of my discussion focused on the elements of our cognitive nature that serve to delimit deliberation. One of these elements is emotion. This is a different level of description, dealing more directly with cognitive function.

There is some overlap here in the argument for the necessity of certain kinds of limitation on evidential considerations, and a more direct discussion about possible ways in which this might occur. I hope in this way to have argued convincingly both that emotion is essential and central to rationality and to have shown how in fact it might take place.

APPENDIX I: WHY REASON?

At the beginning of Chapter II I suggested that one of the meta-questions that we might want to address in a theory of rationality was, "What reasons do we have for requiring reasons, or holding reasons in esteem?" I want to briefly here consider this question.

"Should I follow reason?" ... is a question about whether the game is worth playing. And this is a meaningful question. It might be better to "follow inspiration" than to "follow reason" ... But ... it seems to me obvious that the answer is "Yes, [I should follow reason] because it pays." Deliberation is the only reliable method. Even if there were other reliable methods, we could only tell whether they were reliable by checking them against this method. (Baier, 161-2)

This is how Kurt Baier addresses this question. He suggests that our basis for following reason is ultimately a pragmatic one. We follow reason because it works; it produces results. He continues his discussion and suggests that even if we are going to decide to defer to a higher authority and give up reasoning to follow the will of a leader, that the only way the decision to do this makes any sense is if we reason about our decision to give up reasoning. We might decide, through reasoning, that this individual is better at reasoning for some reason than we are, or that he has worked through some particular set of problems and that we defer to him in the interest of expediency.

Nicholas Rescher reaches much the same conclusion, though with a cautionary note:

The fact of the matter is that we cannot prove that rationality pays - necessarily, or even only over the probabilistic long run. We do not know that acting rationally in the particular case at hand will actually pay off - nor can we even claim with unalloyed assurance that it will probably do so (with real likelihood rather than subjective probability). We can only say that, as best we can judge the matter, it represents the most promising course at our disposal. We have no guarantees - no means are at our disposal for pre-establishing that following

rationality's counsel actually pays. (Rescher,38)

Rescher asserts, rather, that it is in the long run application of rationality to our decision making that we increase our probability of success. "As long as rationality improves the prospects of success, no matter how modestly, its call represents the best bet, the advisable course, the sensible thing to do. No guarantees are necessary." (Rescher,35) Our recourse to rationality as a way of proceeding, Rescher puts forth, is ultimately pragmatic since it is not possible for us to know if it is optimal per se. He also suggests that proceeding rationally offers a means of commonality and mutual understanding for human beings and thus makes communication and co-operation much more readily attainable. "The quest for 'the best available' leads one to fix on that alternative at which others too could be expected to arrive in the circumstances - so that they too can understand one's choices." (Rescher,42) Rescher certainly has a point here though I think it should probably be stated a little more cautiously since, as he himself notes, the addition of any one piece of relevant information may utterly change the determinations of reason, and certainly individuals are often in widely varying epistemic positions. Furthermore, the degree to which individuals reason well will vary widely, and even beyond this there is likely room for varying reasoning (thinking) styles within the domain of rational procedure. I'm sure that he doesn't want to suggest that given the relevant information set there is only one way to proceed. This hints too strongly at an algorithmic process. While it may be true that certain types of reasoning may be carried out algorithmically I don't think that Rescher wants to suggest that this is essential to human reasoning. His point, then, about rationality as a socially binding force

can be appreciated, but the strength of reason as social glue must also be understood to be quite circumscribed.

There is a circularity to the justification of rationality. To proceed in a rational way is what rationality calls for. Rescher suggests, though, that this is exactly how it should be. To seek for reasons for rationality really only exemplifies the nature of rationality. If we are to find a justification for rationality then it will be done so rationally by finding reasons. "But, of course any sensible person not already committed to such a position [of rationality] would want to know if there is any good reason for taking it. And then we are at once back in the sphere of rationality and good reasons." (Rescher,45) Rationality, the looking for reasons, is self-justifying. Once one asks: "Why rationality? Why be rational?", they have already embarked on the search for reasons and rational inquiry. Rescher goes on to note that none of this will satisfy the sceptic. "For him, the lack of guarantees undermines the whole project of rationality." (Rescher,47) We have already seen that for Rescher rationality provides a positive impetus to broaden our knowledge and that when it is considered in this way the sceptic's position is not rationally tenable.

APPENDIX II: CONTEXT AND METHODOLOGY

1. Introduction

In this appendix I discuss some of the broader issues that concern the context and methodology of this thesis. I first discuss the mind-body problem because I see this thesis as fitting into the broader topic of embodiment about which I had originally considered writing. Emotion stands out in the topic of embodiment because it involves essentially both the cognitive and the physiological. In the second section I discuss the problem of free will and determinism both because some of the arguments that de Sousa makes suggest that rationality in the absence of emotion would lead to determinism, and because descriptions of the functioning of mind as mechanical processes often make people suspect that we really are determinate mechanisms. I argue that even if in some sense we are determined, it doesn't really matter. In the last section I talk about the methods of cognitive science and contemporary philosophy of mind which are precisely the things which lead to the fears I have just mentioned. These current models of explanation argue that it is useful to understand the mind as analogous to a computer running software, i.e., as a computational device. There are different types or levels of description that are useful in attempting to understand the functioning of a computational device, and I describe these as the methodological base which underlies my own approach to understanding the human mind.

2 Mind/Body

I want to address here the mind/ body problem because I believe that this is an important aspect of the broader context which the topic of the relation of emotion to rationality fits into because emotion is often conceived of as the aspect of mind that is most closely related to the body. In the western philosophical tradition, especially since Descartes, we have attempted to understand human beings as composed of two distinct elements: body and mind. This description of human beings has tinted, and perhaps tainted, our understanding of the world and ourselves this tradition. These concepts of body and mind still seem clear to us (since we are part of this tradition) but we should not hold to them too strongly.

We typically refer to emotion as an aspect of mind. However, it would seem that in our dualist tradition there are some aspects of mind such as emotion that are often conceived of as being in closer relation to the body than other aspects of mind. Emotion, as a classification, is given both broader and narrower domains of reference, sometimes including, but other times distinct from such things as: passions, moods, inclinations, appetites, and desires.

An example of how emotion can be conceived as an aspect of mind in closer relation to the body can be found in Plato's tripartite division of the soul. Plato divides the soul into the following three parts: 1) the appetitive part; 2) the higher passions or spirited part, and; 3) the rational part. (Plato, 129-138) The first part, involving essentially bodily appetites, is more closely connected to the body than is the rational part at the other end of the spectrum which can act as a counter-force against that aspect of soul connected to the body through appetites. The spirited

part, in Plato's division, which we would tend to associate with what we refer to as emotion, is, it seems to me, mid-point in the relation that holds between mind and body.

3 Free will, Determinism, Compatibilism

In Chapter 1 of *The Rationality of Emotion* de Sousa considers a situation in which I tell a friend of mine something that she has a strongly negative reaction to, namely that I am a homosexual. Later my friend apologizes for her harsh reaction. De Sousa considers which of her reactions I should take to be authentic. He considers the following position, which he calls the chemical view:

Coarsely expressed, the chemical view holds that the first reaction marks nothing more essential than the extent to which this or that hormone or neurotransmitter dominated the organism's chemical balance at the time. As for the second reaction, it is neither more nor less real but simply marks another stage in the friend's chemical history. (de Sousa,13)

De Sousa here seems to give consideration to a kind of chemical determinism. My friend isn't really responsible for her reactions. In important respects they are not even really hers. But can I really view persons I interact with, in this way? Does it even make sense?

Certainly it would seem to be true that our mental states are largely the result of chemical phenomena. My neurons have certain chemical potentials and there are particular causes and effects of their firing. The chemical nature of mind becomes strikingly apparent on certain occasions when we find ourselves acting out of character due to external chemical influences. For example, if I have done mind altering drugs. Even more striking, though, would seem to be the effect of varying levels of hormones. Elevated testosterone certainly seems to be able to make people

aggressive who otherwise would not be. Furthermore, some women seem to experience mood changes corresponding with the changes in their menstrual cycle.

Now the suggestion that de Sousa considers above seems to me to be stronger than this, viz., that we should not consider the mental states of this individual as particularly her mental states, but simply as the manifestations of a chain of chemical events. Certainly insofar as my friend's reaction may have been influenced by some chemical that makes her act in a way that I would consider to be out of character, I would want to take that into consideration; I would judge her to have a reduced level of responsibility for that reaction. Ultimately I would decide whether to dismiss the initial reaction insofar as it proves to be out of character. This holds open the possibility of simply being deceived when my friend decides to moderate her reaction for purely selfish reasons. Needless to say it can be very difficult to be a good judge of character.

The point of view of chemical determinism holds a certain attractiveness insofar as we want to see mental events as being essentially chemical events. The chemical determinist's view may lead to the notion that we should simply dismiss my friend's reaction as a merely chemical event. I think the chemical determinist's view is mistaken because even if we accept this kind of chemical causality, it does not mean that my friend's mental states are not hers, nor even that she is not free and responsible for her own mental dispositions and actions that issue from those mental dispositions.

Any reasonably astute determinist will take account of "apparent" freedom. The astute determinist holds that the individual goes through

quite complicated decision-making processes, experiences emotions, physiological states, etc., which influence these processes, which would lead the individual to think that she is really choosing between genuine alternatives. The causal determinist points out, however, that despite all these things that make it seem like I am free, my various states are causally determined and are quite beyond my control.

Some opponents of the determinist position want to say that despite this kind of description which suggests that we are indeed determined, and thus not free, that there is in fact no good reason to support this position, since there are various possible ways to counter it. One of them is to reject the arguments for determinism on the basis of unacceptable consequences. That is, if we accept determinism it becomes nonsensical to hold individuals responsible for their actions. This would seem to exclude the possibility of both commendation and punishment. It is questionable how a society might be ordered given the absence of these types of actions. There are other attacks aimed more directly at the truth of the determinist position, rather than at its unacceptable consequences, but I shan't go into those right now.

There is a further position with reference to determinism which is referred to as compatibilism. This position does not reject the arguments for the determinist position, but rather suggests that despite the truth of the determinist position human freedom is still possible. It is this position, in defense of which Daniel Dennett argues in *Elbow Room*, that I want to assume.

Consider again de Sousa's example above. He suggests that rather than trying to determine which of my friend's reactions to give more

weight to I could simply move beyond this and chalk all of her reactions up to a determined chain of chemical events. What I am trying to suggest here is that while I can view my friend's reactions as the product of chemical events I cannot, nonetheless, dismiss myself from the need of trying to decide which of her reactions to give more weight to. This is precisely because, despite being the products of chemical events they remain my friend's mental states. Furthermore, I want to maintain, she is free and responsible for these states and the reactions to which they give rise.

4 Modes of Description

In the above section I considered the problem of how I am to decide what kind of weight to give to the different positions my friend takes in reacting to me. Should it be her immediate or her more considered reaction? This led to questions of how I am to view my friend. Is she a free agent or merely a determinate computational system built along chemical lines. The point certainly seems to be moot when we consider that even if the latter is the case we certainly cannot at this point and likely never will be able to give a full account of these processes (at least not with any kind of predictive accuracy). What becomes central here, though, is the question of how I am to view my friend. It really is quite an important methodological question and one that I think Daniel Dennett addresses quite well. The question, then, is, how are we to understand the human mind, and to explain it?

The heirs to cognitive psychology (and to philosophical psychology) beyond behaviourism are various cognitivist accounts of mind that suggest varying accounts of cognition. Much of the theoretical discussion is

premised on a descriptive model of human cognition that sees it, in important respects, as being analogous to the computational functioning of a computer running software.

Daniel Dennett is an important philosophical figure in this tradition and he offers an analysis of an approach we might take toward understanding anything that is like a mind; what he refers to as an intentional system. Dennett suggests that what he has in mind by an intentional system is, in a rough and ready way, "... a system whose behaviour can be - at least sometimes - explained and predicted by relying on ascriptions to the system of beliefs and desires (and hopes, fears, intentions, hunches...)" and that he makes this suggestion "in virtue of the intentionality of the idioms of belief and desire". (Dennett 1978c,3) The notion of intentionality being used here comes from Franz Brentano. It is essentially the idea of object directedness; all thoughts are thoughts of something. Brentano writes:

Every mental phenomenon is characterized by what the scholastics of the Middle Ages called the intentional (and also mental) inexistence of an object, and what we would call, although not in entirely unambiguous terms, the reference to a content, a direction upon an object (by which we are not to understand a reality), or an immanent objectivity. Each one includes something as an object within itself, although not always in the same way. In presentation something is presented, in judgment something is affirmed or denied, in love [something is] loved, in hate [something] is hated, in desire something is desired, etc.

This intentional inexistence is exclusively characteristic of mental phenomena. No physical phenomenon manifests anything similar. Consequently, we can define mental phenomena by saying that they are such phenomena as include an object intentionally within themselves. (Brentano,201)

Brentano suggests here something about the nature of the objects of thought, namely that we can affirm nothing about their reality or existence

from their occurrence as the objects of thought (intentional inexistence), and that it is a defining characteristic of thought that it is intentional.

Now what Dennett has in mind has to do with different ways of describing and coming to an understanding of things that he wants to call intentional systems. Intentionality is the defining characteristic of the mental, according to Brentano, but I will not beg the question by suggesting that anything that we might usefully use intentional terms to describe is a mind.² It is Dennett's concern, though, to suggest different stances for predicting behaviour. This is behaviour in a broad sense, for the different stances will be used in different ways and in different places in order to understand different events in the world, where we may genuinely want to refer to only a certain number as behaviour. Dennett writes:

... a particular thing is an intentional system only in relation to the strategies of someone who is trying to explain and predict its behaviour. ... Consider the case of a chess playing computer, and the different strategies ... one might adopt as its opponent in trying to predict its moves. There are three different stances of interest to us. First there is the design stance. If one knows exactly how the computer is designed (including the impermanent part of its design: its program) one can predict its designed response to any move one makes by following the computation instructions of the program. One's prediction will come true provided only that the computer performs as designed - that is, without breakdown. (Dennett 1978c,4)

Dennett goes on to note that there are different varieties of design stance predictions but they all have in common their reliance on the notion of

² Indeed, intentionality is usually understood as a property of propositions. A proposition is intentional if the truth value of an imbedded clause cannot be determined by the truth value of the whole proposition. For example, if the phrase "John thinks it is raining" is true, it is nonetheless not the case that we can say anything about the truth value of "It is raining".

function, "... which is purpose-relative or teleological. That is, a design of a system breaks it up into larger or smaller functional parts, and design stance predictions are generated assuming that each functional part will function properly." (Dennett 1978c,4) Thus design stance predictions are dependent on the proper functioning of the component functional elements of the object of which we are trying to predict the behaviour.

Physical stance prediction involves taking into account the physical state of the object of which we are trying to predict the behaviour. Considering what we know about the physical state of the object and what we know about the laws of nature we can make a prediction of behaviour or events. There are two types of situations where we typically use physical stance predictions. One is in situations where we are considering fairly simple physical events and the physical stance is the only possible way predictive stance. For example, I note that the limb on this tree is cracked. Furthermore, I note that we are getting freezing rain and that ice is accumulating on the branch. Now I know that while the branch in its cracked state is strong enough to support its own weight I can try to estimate to what degree it is weakened and how much weight the ice will add and consider whether this will cause the branch to break and fall. The other type of situation where we would typically use a physical stance description is that where we would normally use a design stance method but for some reason there is a malfunction and the design stance no longer works. Using a very simple example, if I turn on my computer and nothing happens, unlike what I would expect from design, I immediately consider possible physical causes. First I make sure it is plugged in. Then I make sure that the fuse is not blown. I continue along a line of

physical investigation until I discover what has gone wrong.

Dennett in his example of a chess playing computer notes:

The best chess playing computers these days are practically inaccessible to prediction from either the design stance or the physical stance; they have become too complex for even their own designers to view from the design stance. A man's best hope of defeating such a machine in a chess match is to predict its responses by figuring out as best he can what the best or most rational move would be, given the rules and goals of chess. That is, one assumes not only (1) that the machine will function as designed, but (2) that the design is optimal as well, that the computer will "choose" the most rational move. (Dennett 1978c,5)

Dennett goes on to note that the predictions will likely fail if either of our two assumptions proves to be unwarranted. And we are nonetheless left, when trying to make predictions in this manner, with all the indeterminacy that we face when trying to make predictions about the actions of other intelligent beings, for this is essentially how we are treating the computer in this instance. The assumption of rationality involved in this model of prediction is only an assumption of optimal design with respect to certain goals (in this instance those of winning pieces and ultimately the game), and restraints (the rules of chess).

The intentional stance prediction, then, ascribes beliefs and desires to the machine (or whatever else we may be trying to predict the behaviour of) and considers what it would be rational to do given those beliefs and desires. The ascriptions of beliefs and desires are in terms of intentional idioms. Beliefs are beliefs about something and desires are desires for something. The computer presents a very limited case of the use of intentional stance predictions insofar as I essentially know what the computer believes and desires. Its beliefs are the positions of the pieces on the board. Its desire is to win the game. Unlike a human opponent I

needn't consider whether the computer is growing tired of the game and considering letting me win to get it over with, or simply throwing in the towel. Thus, I might reason as follows. My queen is in a defensive position protecting my king, and is a very valuable piece in its own right. So, if given the opportunity to capture my queen without the loss of an important piece the computer will do so. But, the computer is in a position to do that now. Therefore, the computer is about to capture my queen. This is essentially how intentional stance explanations function, though they can certainly be much more involved than this.

Intentional stance explanations are the real basis of folk psychology - what people typically use in their everyday lives to predict the behaviour of others. This is not to say that intentional stance explanation has not been advanced by psychologists in ways not employed in everyday life. Indeed importing of pieces of a design stance approach in order to flesh out certain things is often done. For instance, if I am aware of research that suggests that people tend to persevere in believing things even after their original reasons for coming to those beliefs have been refuted, then I will take that into consideration when I try to predict how someone will behave. That is, according to a straightforward folk psychological account we might expect someone to give up a belief in this circumstance. But armed with evidence to the contrary I expect that she may very well keep such a belief unless she is given stronger reasons for abandoning it than the mere fact that her original reasons for coming to believe it have been refuted.

The intentional stance model of explaining mind underlies much of what I say about emotion and rationality.

Bibliography

- Alston, William P. 1967. "Emotion and Feeling." In *The Encyclopedia of Philosophy*, ed. Paul Edwards. New York and London: Macmillan.
- Aristotle. 1990. *The Nichomachean Ethics*, trans., Sir David Ross, revision J.L. Ackrill & J.O. Urmson. Oxford: Oxford University Press.
- Baier, Kurt. 1965. *The Moral Point of View*. New York: Random House.
- Belenky, Mary Field, et. al. 1986. *Women's Ways of Knowing*. New York: Basic Books Inc.
- Brentano, Franz. 1874. *Psychologie vom empirischen Standpunkt*, Vol.I, bk.II, Ch.1, as cited in Chisholm, Roderick. 1963.
- Calhoun, Cheshire, and Robert C. Solomon, eds. 1984. *What is an Emotion? Classical Readings in Philosophical Psychology*. New York and Oxford: Oxford University Press.
- Chisholm, Roderick. 1963. "Intentionality" In *The Encyclopedia of Philosophy*, ed. Paul Edwards. New York and London: Macmillan
- Code, Lorraine. 1987. *Epistemic Responsibility*. Hanover New Hampshire: University Press of New England.
- Davidson, Donald. 1980. "Actions, Reasons and Causes." In *Essays on Actions and Events*. Oxford: Clarendon Press.
- Dennett, Daniel C. 1978a. *Brainstorms*. Cambridge Massachusetts: MIT Press/Bradford Books,
- Dennett, Daniel C. 1978b. "On Giving Libertarians What They Say They Want." In Dennett 1978a.
- Dennett, Daniel C. 1978c. "Intentional Systems." In Dennett 1978a
- Dennett, Daniel C. 1984. *Elbow Room*. Cambridge Massachusetts: MIT Press/Bradford Books.
- de Sousa, Ronald, 1987. *The Rationality of Emotion*. Cambridge Massachusetts: MIT Press.
- Fodor, Jerry A. 1983. *The Modularity of Mind*. Cambridge Massachusetts: MIT Press.
- Fodor, Jerry A. 1987. "Modules, Frames, Fridgeons, Sleeping Dogs and the Music of the Spheres." In Pylyshyn 1987.
- Garaets, Theodore F., ed. 1979. *Rationality Today/La Rationalite Aujourd'hui*. Ottawa: The University of Ottawa Press.

- Haugeland, John. 1985. *Artificial Intelligence: The Very Idea*. Cambridge Massachusetts: MIT Press.
- Harman, Gilbert. 1986. *Change in View: Principles of Reasoning*. Cambridge Massachusetts: Bradford Books/ MIT Press.
- Habermas, Jurgen. 1971. *Knowledge and Human Interests*. trans. Shapiro, Jeremy J. Boston: Beacon Press.
- Hume, David. 1888. *A Treatise on Human Nature*. ed. L.A. Selby-Bigge. Oxford: Oxford/Clarendon.
- Irani, K.D., G.E. Myers, eds. 1983. *Emotions: Philosophical Studies*. New York: Haven Publications.
- Irani, K.D. 1983. "Understanding Emotion." In Irani, K.D., G.E. Myers, eds. 1983.
- James, William. 1971. *A William James Reader*, ed. Gay Wilson Allen. Boston: Houghton Mifflin Company.
- James, William. 1980a. "The Will to Believe." In Thayer 1980.
- James, William. 1980b. "What Pragmatism Means." In Thayer 1980.
- Kant, Immanuel. 1956, *The Groundwork of the Metaphysic of Morals*, trans. H.J. Paton. New York: Harper & Row, by arrangement with Hutchinson & Co. Ltd., London.
- Kuhn, Thomas S. 1969. *The Structure of Scientific Revolutions*. 2nd ed. Chicago: Chicago University Press.
- Morick, Harold. 1980. *Challenges to Empiricism*. Indianapolis Indiana: Hackett Publishing Co.
- Plato. 1981. *The Republic*, trans. Cornford, Francis MacDonald. London: Oxford University Press.
- Pylyshyn, Zenon W., ed. 1987. *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*. Norwood, New Jersey: Ablex Publishing Corp.
- Quine, W.V.O. 1953. "Two Dogma's of Empiricism" in *From a Logical Point of View*. 1953. Cambridge Massachusetts: Harvard University Press.
- Rescher, Nicholas. 1988. *Rationality*. Oxford: Clarendon Press.
- Rorty, Amelie Oksenberg, ed. 1980. *Explaining Emotions*. Berkeley: Univeresity of California Press.
- Rorty, Amelie Oksenberg, ed. 1980. *The Identities of Persons*. Berkeley: Univeresity of California Press.

Rorty, Richard. 1979. *Philosophy and the Mirror of Nature*. Princeton, New Jersey: Princeton University Press.

Sartre, Jean Paul. 1948. *The Emotions: Outline of a Theory*. trans. Bernard Frechtman. New York: Philosophical Library.

Sellars, Wilfred, 1980. "Empiricism and the Philosophy of Mind." In Morick 1980.

Siegel, Harvey. 1988. *Educating Reason*. New York: Routledge.

Solomon, Robert C. 1976. *The Passions*. Garden City, New York: Anchor Press/Doubleday,

Szabados, Bela. 1979. "A Note on Irrationality" In Garaets 1979.

Tamny, Martin and K.D. Irani, eds. 1986. *Rationality in thought and Action*. New York: Greenwood Press.

Thayer, H.S., ed. 1982. *Pragmatism: The Classic Writings*. Indianapolis Indiana: Hackett Publishing Co.

Wadia, Pheroze. 1986. "Reasoning, Believing and Willing." In Tamny and Irani, eds., 1986.

Wittgenstein, Ludwig. 1958. *Philosophical Investigations*. Oxford: Basil Blackwell.

VITA AUCTORIS

John Mckay was born in Windsor, Ontario in 1965. He entered the University of Windsor in September of 1983 and recieved a General B.A. in 1987. He received the Honours B.A. in Philosophy in 1990. He hopes to recieve the M.A. in Philosophy in 1993 from the University of Windsor and is also currently pursuing the M.A. in Spanish from Middlebury College, Vermont.